SEQUENCE LISTING

<110> BASF Plant Science GmbH

       SweTree Technologies AB

<120> POST HARVEST CONTROL OF GENETICALLY MODIFIED CROP GROWTH EM-
PLOYING D-AMINO ACID COMPOUNDS

<130> PF 55447 EP

<160> 16

<170> PatentIn version 3.1

<210> 1
<211> 1160
<212> DNA
<213> Rhodosporidium toruloides

<220>
<221> CDS
<222> (1)..(1104)
<223> coding for DAAO

<400> 1

```
atg cac tcg cag aag cgc gtc gtt gtc ctc gga tca ggc gtt atc ggt      48
Met His Ser Gln Lys Arg Val Val Val Leu Gly Ser Gly Val Ile Gly
1               5                   10                  15
ctg agc agc gcc ctc atc ctc gct cgg aag ggc tac agc gtg cat att      96
Leu Ser Ser Ala Leu Ile Leu Ala Arg Lys Gly Tyr Ser Val His Ile
                20                  25                  30
ctc gcg cgc gac ttg ccg gag gac gtc tcg agc cag act ttc gct tca     144
Leu Ala Arg Asp Leu Pro Glu Asp Val Ser Ser Gln Thr Phe Ala Ser
            35                  40                  45
cca tgg gct ggc gcg aat tgg acg cct ttc atg acg ctt aca gac ggt     192
Pro Trp Ala Gly Ala Asn Trp Thr Pro Phe Met Thr Leu Thr Asp Gly
        50                  55                  60
cct cga caa gca aaa tgg gaa gaa tcg act ttc aag aag tgg gtc gag     240
Pro Arg Gln Ala Lys Trp Glu Glu Ser Thr Phe Lys Lys Trp Val Glu
65                  70                  75                  80
ttg gtc ccg acg ggc cat gcc atg tgg ctc aag ggg acg agg cgg ttc     288
Leu Val Pro Thr Gly His Ala Met Trp Leu Lys Gly Thr Arg Arg Phe
                85                  90                  95
```

```
      gcg cag aac gaa gac ggc ttg ctc ggg cac tgg tac aag gac atc acg     336
      Ala Gln Asn Glu Asp Gly Leu Leu Gly His Trp Tyr Lys Asp Ile Thr
                  100             105                 110
      cca aat tac cgc ccc ctc cca tct tcc gaa tgt cca cct ggc gct atc     384
 5    Pro Asn Tyr Arg Pro Leu Pro Ser Ser Glu Cys Pro Pro Gly Ala Ile
                  115             120                 125
      ggc gta acc tac gac acc ctc tcc gtc cac gca cca aag tac tgc cag     432
      Gly Val Thr Tyr Asp Thr Leu Ser Val His Ala Pro Lys Tyr Cys Gln
                  130             135                 140
10    tac ctt gca aga gag ctg cag aag ctc ggc gcg acg ttt gag aga cgg     480
      Tyr Leu Ala Arg Glu Leu Gln Lys Leu Gly Ala Thr Phe Glu Arg Arg
      145                 150             155                 160
      acc gtt acg tcg ctt gag cag gcg ttc gac ggt gcg gat ttg gtg gtc     528
      Thr Val Thr Ser Leu Glu Gln Ala Phe Asp Gly Ala Asp Leu Val Val
15                        165             170                 175
      aac gct acg gga ctt ggc gcc aag tcg att gcg ggc atc gac gac caa     576
      Asn Ala Thr Gly Leu Gly Ala Lys Ser Ile Ala Gly Ile Asp Asp Gln
                  180             185                 190
      gcc gcc gag cca atc cgc ggg caa acc gtc ctc gtc aag tcc cca tgc     624
20    Ala Ala Glu Pro Ile Arg Gly Gln Thr Val Leu Val Lys Ser Pro Cys
                  195             200                 205
      aag cga tgc acg atg gac tcg tcc gac ccc gct tct ccc gcc tac atc     672
      Lys Arg Cys Thr Met Asp Ser Ser Asp Pro Ala Ser Pro Ala Tyr Ile
                  210             215                 220
25    att ccc cga cca ggt ggc gaa gtc atc tgc ggc ggg acg tac ggc gtg     720
      Ile Pro Arg Pro Gly Gly Glu Val Ile Cys Gly Gly Thr Tyr Gly Val
      225                 230             235                 240
      gga gac tgg gac ttg tct gtc aac cca gag acg gtc cag cgg atc ctc     768
      Gly Asp Trp Asp Leu Ser Val Asn Pro Glu Thr Val Gln Arg Ile Leu
30                        245             250                 255
      aag cac tgc ttg cgc ctc gac ccg acc atc tcg agc gac gga acg atc     816
      Lys His Cys Leu Arg Leu Asp Pro Thr Ile Ser Ser Asp Gly Thr Ile
                  260             265                 270
      gaa ggc atc gag gtc ctc cgc cac aac gtc ggc ttg cga cct gca cga     864
35    Glu Gly Ile Glu Val Leu Arg His Asn Val Gly Leu Arg Pro Ala Arg
                  275             280                 285
      cga ggc gga ccc cgc gtt gag gca gaa cgg atc gtc ctg cct ctc gac     912
      Arg Gly Gly Pro Arg Val Glu Ala Glu Arg Ile Val Leu Pro Leu Asp
                  290             295                 300
40    cgg aca aag tcg ccc ctc tcg ctc ggc agg ggc agc gca cga gcg gcg     960
      Arg Thr Lys Ser Pro Leu Ser Leu Gly Arg Gly Ser Ala Arg Ala Ala
      305                 310             315                 320
      aag gag aag gag gtc acg ctt gtg cat gcg tat ggc ttc tcg agt gcg    1008
      Lys Glu Lys Glu Val Thr Leu Val His Ala Tyr Gly Phe Ser Ser Ala
45                        325             330                 335
      gga tac cag cag agt tgg ggc gcg gcg gag gat gtc gcg cag ctc gtc    1056
      Gly Tyr Gln Gln Ser Trp Gly Ala Ala Glu Asp Val Ala Gln Leu Val
```

```
              340                    345                    350
    gac gag gcg ttc cag cgg tac cac ggc gcg gcg cgg gag tcg aag ttg      1104
    Asp Glu Ala Phe Gln Arg Tyr His Gly Ala Ala Arg Glu Ser Lys Leu
              355                    360                    365
    tagggcggga tttgtggctg tattgcgggc atctacaaga aaaaaaaaaa aaaaa        1160
```

```
<210>  2
<211>  368
<212>  PRT
<213>  Rhodosporidium toruloides
```

```
<400>  2
Met His Ser Gln Lys Arg Val Val Val Leu Gly Ser Gly Val Ile Gly
1               5                   10                  15

Leu Ser Ser Ala Leu Ile Leu Ala Arg Lys Gly Tyr Ser Val His Ile
            20                  25                  30

Leu Ala Arg Asp Leu Pro Glu Asp Val Ser Ser Gln Thr Phe Ala Ser
        35                  40                  45

Pro Trp Ala Gly Ala Asn Trp Thr Pro Phe Met Thr Leu Thr Asp Gly
    50                  55                  60

Pro Arg Gln Ala Lys Trp Glu Glu Ser Thr Phe Lys Lys Trp Val Glu
65                  70                  75                  80

Leu Val Pro Thr Gly His Ala Met Trp Leu Lys Gly Thr Arg Arg Phe
                85                  90                  95

Ala Gln Asn Glu Asp Gly Leu Leu Gly His Trp Tyr Lys Asp Ile Thr
            100                 105                 110

Pro Asn Tyr Arg Pro Leu Pro Ser Ser Glu Cys Pro Pro Gly Ala Ile
            115                 120                 125

Gly Val Thr Tyr Asp Thr Leu Ser Val His Ala Pro Lys Tyr Cys Gln
        130                 135                 140

Tyr Leu Ala Arg Glu Leu Gln Lys Leu Gly Ala Thr Phe Glu Arg Arg
145                 150                 155                 160

Thr Val Thr Ser Leu Glu Gln Ala Phe Asp Gly Ala Asp Leu Val Val
                165                 170                 175

Asn Ala Thr Gly Leu Gly Ala Lys Ser Ile Ala Gly Ile Asp Asp Gln
            180                 185                 190
```

```
     Ala Ala Glu Pro Ile Arg Gly Gln Thr Val Leu Val Lys Ser Pro Cys
             195                 200                 205

5    Lys Arg Cys Thr Met Asp Ser Ser Asp Pro Ala Ser Pro Ala Tyr Ile
             210                 215                 220

     Ile Pro Arg Pro Gly Gly Glu Val Ile Cys Gly Gly Thr Tyr Gly Val
     225                 230                 235                 240
10
     Gly Asp Trp Asp Leu Ser Val Asn Pro Glu Thr Val Gln Arg Ile Leu
                     245                 250                 255

     Lys His Cys Leu Arg Leu Asp Pro Thr Ile Ser Ser Asp Gly Thr Ile
15           260                 265                 270

     Glu Gly Ile Glu Val Leu Arg His Asn Val Gly Leu Arg Pro Ala Arg
             275                 280                 285

20   Arg Gly Gly Pro Arg Val Glu Ala Glu Arg Ile Val Leu Pro Leu Asp
             290                 295                 300

     Arg Thr Lys Ser Pro Leu Ser Leu Gly Arg Gly Ser Ala Arg Ala Ala
     305                 310                 315                 320
25
     Lys Glu Lys Glu Val Thr Leu Val His Ala Tyr Gly Phe Ser Ser Ala
                     325                 330                 335

     Gly Tyr Gln Gln Ser Trp Gly Ala Ala Glu Asp Val Ala Gln Leu Val
30                   340                 345                 350

     Asp Glu Ala Phe Gln Arg Tyr His Gly Ala Ala Arg Glu Ser Lys Leu
             355                 360                 365

35   <210>   3
     <211>   1005
     <212>   DNA
     <213>   Caenorhabditis elegans

40   <220>
     <221>   CDS
     <222>   (1)..(1002)
     <223>   coding for DAAO

45   <400>   3
     atg gca aac ata att ccg aag att gca att atc ggc gaa gga gtc att        48
     Met Ala Asn Ile Ile Pro Lys Ile Ala Ile Ile Gly Glu Gly Val Ile
```

```
        1               5                       10                      15
        gga tgt act tca gca ctt caa ata tca aaa gct ata cca aat gcg aaa        96
        Gly Cys Thr Ser Ala Leu Gln Ile Ser Lys Ala Ile Pro Asn Ala Lys
                        20                      25                      30
 5      ata act gtg ctc cac gat aaa cca ttt aaa aaa tcg tgc agt gca gga       144
        Ile Thr Val Leu His Asp Lys Pro Phe Lys Lys Ser Cys Ser Ala Gly
                    35                      40                      45
        cca gca gga tta ttt aga atc gat tat gag gag aat act gaa tac gga       192
        Pro Ala Gly Leu Phe Arg Ile Asp Tyr Glu Glu Asn Thr Glu Tyr Gly
10          50                      55                      60
        cgt gct tct ttc gcc tgg ttc tca cat ctc tat cgc act aca aaa gga       240
        Arg Ala Ser Phe Ala Trp Phe Ser His Leu Tyr Arg Thr Thr Lys Gly
        65                      70                      75                      80
        tcc gaa acc ggc gtg aaa tta gtt tct gga cat att caa tcc gac aac       288
15      Ser Glu Thr Gly Val Lys Leu Val Ser Gly His Ile Gln Ser Asp Asn
                        85                      90                      95
        ttg gag tca ttg aag caa caa caa aga gcc tat ggc gat att gtg tac       336
        Leu Glu Ser Leu Lys Gln Gln Gln Arg Ala Tyr Gly Asp Ile Val Tyr
                    100                     105                     110
20      aac ttt aga ttc ttg gat gat aga gaa cgg ctg gac att ttt ccc gaa       384
        Asn Phe Arg Phe Leu Asp Asp Arg Glu Arg Leu Asp Ile Phe Pro Glu
                    115                     120                     125
        cca tca aag cac tgc att cac tac acc gcc tac gca tca gaa ggt aac       432
        Pro Ser Lys His Cys Ile His Tyr Thr Ala Tyr Ala Ser Glu Gly Asn
25          130                     135                     140
        aag tac gtg cct tat ttg aag aat ttg ctg ctt gag caa aaa atc gag       480
        Lys Tyr Val Pro Tyr Leu Lys Asn Leu Leu Leu Glu Gln Lys Ile Glu
        145                     150                     155                     160
        ttc aag caa caa gaa gtg acg agt ttg gac gca gtc gcc gac gct ggt       528
30      Phe Lys Gln Gln Glu Val Thr Ser Leu Asp Ala Val Ala Asp Ala Gly
                        165                     170                     175
        tac gat gtt att gta aac tgc gca ggc ttg tac ggt gga aag ttg gct       576
        Tyr Asp Val Ile Val Asn Cys Ala Gly Leu Tyr Gly Gly Lys Leu Ala
                    180                     185                     190
35      ggt gat gac gat act tgc tac ccc att aga gga gtc att ttg gaa gtt       624
        Gly Asp Asp Asp Thr Cys Tyr Pro Ile Arg Gly Val Ile Leu Glu Val
                    195                     200                     205
        gat gca cca tgg cac aag cac ttc aat tat cga gac ttt act act ttc       672
        Asp Ala Pro Trp His Lys His Phe Asn Tyr Arg Asp Phe Thr Thr Phe
40          210                     215                     220
        aca att cca aaa gag cac agc gtg gtg gtt ggg tcc acc aag cag gac       720
        Thr Ile Pro Lys Glu His Ser Val Val Val Gly Ser Thr Lys Gln Asp
        225                     230                     235                     240
        aat cga tgg gat ttg gag atc acc gac gag gat aga aat gat att ttg       768
45      Asn Arg Trp Asp Leu Glu Ile Thr Asp Glu Asp Arg Asn Asp Ile Leu
                        245                     250                     255
        aaa cga tac att gct tta cat cct gga atg aga gag cca aag att atc       816
```

```
        Lys Arg Tyr Ile Ala Leu His Pro Gly Met Arg Glu Pro Lys Ile Ile
                    260                 265                 270
        aaa gaa tgg tca gca ctt cgc ccg gga cgt aag cat gtc aga att gaa      864
        Lys Glu Trp Ser Ala Leu Arg Pro Gly Arg Lys His Val Arg Ile Glu
  5                 275                 280                 285
        gcg cag aag agg aca tct gtt gga aac tca aaa gat tat atg gtt gtg      912
        Ala Gln Lys Arg Thr Ser Val Gly Asn Ser Lys Asp Tyr Met Val Val
                    290                 295                 300
        cat cac tat ggt cac ggg agc aac gga ttc acg ttg ggt tgg gga aca      960
 10     His His Tyr Gly His Gly Ser Asn Gly Phe Thr Leu Gly Trp Gly Thr
        305                 310                 315                 320
        gca att gaa gca act aaa ctt gtt aag act gca cta gga tta taa        1005
        Ala Ile Glu Ala Thr Lys Leu Val Lys Thr Ala Leu Gly Leu
                    325                 330
 15
        <210>   4
        <211>   334
        <212>   PRT
        <213>   Caenorhabditis elegans
 20
        <400>   4
        Met Ala Asn Ile Ile Pro Lys Ile Ala Ile Ile Gly Glu Gly Val Ile
        1                 5                   10                  15
 25     Gly Cys Thr Ser Ala Leu Gln Ile Ser Lys Ala Ile Pro Asn Ala Lys
                        20                  25                  30
        Ile Thr Val Leu His Asp Lys Pro Phe Lys Lys Ser Cys Ser Ala Gly
                        35                  40                  45
 30
        Pro Ala Gly Leu Phe Arg Ile Asp Tyr Glu Glu Asn Thr Glu Tyr Gly
                    50                  55                  60
        Arg Ala Ser Phe Ala Trp Phe Ser His Leu Tyr Arg Thr Thr Lys Gly
 35     65                  70                  75                  80
        Ser Glu Thr Gly Val Lys Leu Val Ser Gly His Ile Gln Ser Asp Asn
                        85                  90                  95
 40     Leu Glu Ser Leu Lys Gln Gln Gln Arg Ala Tyr Gly Asp Ile Val Tyr
                    100                 105                 110
        Asn Phe Arg Phe Leu Asp Asp Arg Glu Arg Leu Asp Ile Phe Pro Glu
                    115                 120                 125
 45
        Pro Ser Lys His Cys Ile His Tyr Thr Ala Tyr Ala Ser Glu Gly Asn
                    130                 135                 140
```

```
      Lys Tyr Val Pro Tyr Leu Lys Asn Leu Leu Leu Glu Gln Lys Ile Glu
      145                 150                 155                 160

   5  Phe Lys Gln Gln Glu Val Thr Ser Leu Asp Ala Val Ala Asp Ala Gly
                          165                 170                 175

      Tyr Asp Val Ile Val Asn Cys Ala Gly Leu Tyr Gly Gly Lys Leu Ala
                      180                 185                 190
  10
      Gly Asp Asp Asp Thr Cys Tyr Pro Ile Arg Gly Val Ile Leu Glu Val
                  195                 200                 205

      Asp Ala Pro Trp His Lys His Phe Asn Tyr Arg Asp Phe Thr Thr Phe
  15      210                 215                 220

      Thr Ile Pro Lys Glu His Ser Val Val Val Gly Ser Thr Lys Gln Asp
      225                 230                 235                 240

  20  Asn Arg Trp Asp Leu Glu Ile Thr Asp Glu Asp Arg Asn Asp Ile Leu
                      245                 250                 255

      Lys Arg Tyr Ile Ala Leu His Pro Gly Met Arg Glu Pro Lys Ile Ile
                  260                 265                 270
  25
      Lys Glu Trp Ser Ala Leu Arg Pro Gly Arg Lys His Val Arg Ile Glu
                  275                 280                 285

      Ala Gln Lys Arg Thr Ser Val Gly Asn Ser Lys Asp Tyr Met Val Val
  30      290                 295                 300

      His His Tyr Gly His Gly Ser Asn Gly Phe Thr Leu Gly Trp Gly Thr
      305                 310                 315                 320

  35  Ala Ile Glu Ala Thr Lys Leu Val Lys Thr Ala Leu Gly Leu
                      325                 330
```

```
      <210>  5
      <211>  1186
  40  <212>  DNA
      <213>  Nectria haematococca

      <220>
      <221>  CDS
  45  <222>  (42)..(1124)
      <223>  coding for DAAO
```

```
<400>  5
agcgacttga atttagcgaa aagaacttgt caaccacaat c atg tcc aac aca atc      56
                                                Met Ser Asn Thr Ile
                                                 1                 5
gtc gtc gtt ggt gcc ggt gtc att ggc ttg acg tcg gcc ttg ttg ctc      104
Val Val Val Gly Ala Gly Val Ile Gly Leu Thr Ser Ala Leu Leu Leu
                10                  15                  20
tcc aag aac aag ggc aac aag atc acc gtc gtg gcc aag cac atg ccc      152
Ser Lys Asn Lys Gly Asn Lys Ile Thr Val Val Ala Lys His Met Pro
            25                  30                  35
ggc gac tat gac gtt gaa tac gcc tcg cct ttt gct ggt gcc aac cac      200
Gly Asp Tyr Asp Val Glu Tyr Ala Ser Pro Phe Ala Gly Ala Asn His
        40                  45                  50
tcc ccc atg gcg acg gaa gag agc agc gaa tgg gaa cgt cgc act tgg      248
Ser Pro Met Ala Thr Glu Glu Ser Ser Glu Trp Glu Arg Arg Thr Trp
        55                  60                  65
tac gag ttt aag aga ctg gtc gag gag gtc cct gag gcc ggt gtt cat      296
Tyr Glu Phe Lys Arg Leu Val Glu Glu Val Pro Glu Ala Gly Val His
70                  75                  80                  85
ttc cag aag tct cgc atc cag agg cgc aat gtg gac act gaa aag gcg      344
Phe Gln Lys Ser Arg Ile Gln Arg Arg Asn Val Asp Thr Glu Lys Ala
                90                  95                  100
cag agg tct ggt ttc cca gac gcc ctc ttc tcg aaa gaa ccc tgg ttc      392
Gln Arg Ser Gly Phe Pro Asp Ala Leu Phe Ser Lys Glu Pro Trp Phe
            105                 110                 115
aag aac atg ttt gag gac ttc cgt gag cag cac cct agc gag gtc atc      440
Lys Asn Met Phe Glu Asp Phe Arg Glu Gln His Pro Ser Glu Val Ile
            120                 125                 130
ccc ggt tac gac tct ggc tgc gag ttc aca tcg gtg tgc atc aac acg      488
Pro Gly Tyr Asp Ser Gly Cys Glu Phe Thr Ser Val Cys Ile Asn Thr
            135                 140                 145
gcc atc tac ctc ccc tgg ctc ctc ggc cag tgc atc aag aat ggc gtc      536
Ala Ile Tyr Leu Pro Trp Leu Leu Gly Gln Cys Ile Lys Asn Gly Val
150                 155                 160                 165
atc gtc aag cgc gcc atc ctc aac gac att agc gag gcc aag aag ctg      584
Ile Val Lys Arg Ala Ile Leu Asn Asp Ile Ser Glu Ala Lys Lys Leu
                170                 175                 180
agc cac gcg ggc aag acg ccc aat atc atc gtc aac gcc acg ggt ctc      632
Ser His Ala Gly Lys Thr Pro Asn Ile Ile Val Asn Ala Thr Gly Leu
            185                 190                 195
ggc tcc tac aag ctg ggc ggt gtc gag gac aag acc atg gcg cct gcg      680
Gly Ser Tyr Lys Leu Gly Gly Val Glu Asp Lys Thr Met Ala Pro Ala
            200                 205                 210
cgg gga cag att gtg gtt gtg cgc aac gag agc agc ccc atg ctc ctc      728
Arg Gly Gln Ile Val Val Val Arg Asn Glu Ser Ser Pro Met Leu Leu
```

```
          215                    220                    225
      act tca ggt gtc gag gac ggc ggt gct gat gtc atg tac ttg atg cag      776
      Thr Ser Gly Val Glu Asp Gly Gly Ala Asp Val Met Tyr Leu Met Gln
      230                    235                    240                    245
  5   cga gca gct ggc ggt ggc acc atc ctg ggc ggt acc tac gac gtt ggc      824
      Arg Ala Ala Gly Gly Gly Thr Ile Leu Gly Gly Thr Tyr Asp Val Gly
                        250                    255                    260
      aac tgg gag tct cag cca gac ccc aac atc gcg aat cgc atc atg cag      872
      Asn Trp Glu Ser Gln Pro Asp Pro Asn Ile Ala Asn Arg Ile Met Gln
 10                     265                    270                    275
      cgc atc gtc gag gtg cgg ccc gag att gcc aac ggc aag ggc gtc aag      920
      Arg Ile Val Glu Val Arg Pro Glu Ile Ala Asn Gly Lys Gly Val Lys
                        280                    285                    290
      ggg ctg agc gtg atc cga cac gcc gtc ggc atg cgg ccg tgg cga aag      968
 15   Gly Leu Ser Val Ile Arg His Ala Val Gly Met Arg Pro Trp Arg Lys
          295                    300                    305
      gac gga gtc agg atc gag gag gag aag ctg gat gat gag act tgg atc     1016
      Asp Gly Val Arg Ile Glu Glu Glu Lys Leu Asp Asp Glu Thr Trp Ile
      310                    315                    320                    325
 20   gtg cac aac tac gga cac tct gga tgg ggt tac cag ggt tcg tat ggt     1064
      Val His Asn Tyr Gly His Ser Gly Trp Gly Tyr Gln Gly Ser Tyr Gly
                        330                    335                    340
      tgt gct gag aat gta gtc cag ttg gtt gac aag gtc ggc aag gcg gcc     1112
      Cys Ala Glu Asn Val Val Gln Leu Val Asp Lys Val Gly Lys Ala Ala
 25                     345                    350                    355
      aag tct aag ctg tagttgaaaa ggcctgaatg agtaatagta attggatatt         1164
      Lys Ser Lys Leu
                  360
      ggaaataccg tatttgccct cg                                            1186
 30
      <210>  6
      <211>  361
      <212>  PRT
      <213>  Nectria haematococca
 35
      <400>  6
      Met Ser Asn Thr Ile Val Val Val Gly Ala Gly Val Ile Gly Leu Thr
      1              5                    10                   15

 40   Ser Ala Leu Leu Leu Ser Lys Asn Lys Gly Asn Lys Ile Thr Val Val
                  20                   25                   30

      Ala Lys His Met Pro Gly Asp Tyr Asp Val Glu Tyr Ala Ser Pro Phe
                  35                   40                   45
 45
      Ala Gly Ala Asn His Ser Pro Met Ala Thr Glu Glu Ser Ser Glu Trp
          50                   55                   60
```

```
      Glu Arg Arg Thr Trp Tyr Glu Phe Lys Arg Leu Val Glu Glu Val Pro
      65              70              75              80

 5    Glu Ala Gly Val His Phe Gln Lys Ser Arg Ile Gln Arg Arg Asn Val
                      85              90              95

      Asp Thr Glu Lys Ala Gln Arg Ser Gly Phe Pro Asp Ala Leu Phe Ser
                  100             105             110
10
      Lys Glu Pro Trp Phe Lys Asn Met Phe Glu Asp Phe Arg Glu Gln His
              115             120             125

      Pro Ser Glu Val Ile Pro Gly Tyr Asp Ser Gly Cys Glu Phe Thr Ser
15            130             135             140

      Val Cys Ile Asn Thr Ala Ile Tyr Leu Pro Trp Leu Leu Gly Gln Cys
      145             150             155             160

20    Ile Lys Asn Gly Val Ile Val Lys Arg Ala Ile Leu Asn Asp Ile Ser
                      165             170                 175

      Glu Ala Lys Lys Leu Ser His Ala Gly Lys Thr Pro Asn Ile Ile Val
                  180             185             190
25
      Asn Ala Thr Gly Leu Gly Ser Tyr Lys Leu Gly Gly Val Glu Asp Lys
                  195             200             205

      Thr Met Ala Pro Ala Arg Gly Gln Ile Val Val Val Arg Asn Glu Ser
30            210             215             220

      Ser Pro Met Leu Leu Thr Ser Gly Val Glu Asp Gly Gly Ala Asp Val
      225             230             235             240

35    Met Tyr Leu Met Gln Arg Ala Ala Gly Gly Gly Thr Ile Leu Gly Gly
                      245             250             255

      Thr Tyr Asp Val Gly Asn Trp Glu Ser Gln Pro Asp Pro Asn Ile Ala
                  260             265             270
40
      Asn Arg Ile Met Gln Arg Ile Val Glu Val Arg Pro Glu Ile Ala Asn
                  275             280             285

      Gly Lys Gly Val Lys Gly Leu Ser Val Ile Arg His Ala Val Gly Met
45            290             295             300

      Arg Pro Trp Arg Lys Asp Gly Val Arg Ile Glu Glu Glu Lys Leu Asp
```

```
         305                    310                  315                    320


         Asp Glu Thr Trp Ile Val His Asn Tyr Gly His Ser Gly Trp Gly Tyr
                             325                  330                    335
 5

         Gln Gly Ser Tyr Gly Cys Ala Glu Asn Val Val Gln Leu Val Asp Lys
                     340                  345                  350


         Val Gly Lys Ala Ala Lys Ser Lys Leu
10                   355                  360


         <210>   7
         <211>   1071
         <212>   DNA
15       <213>   Trigonopsis variabilis


         <220>
         <221>   CDS
         <222>   (1)..(1068)
20       <223>


         <400>  7
         atg gct aaa atc gtt gtt att ggt gcc ggt gtt gcc ggt tta act aca       48
         Met Ala Lys Ile Val Val Ile Gly Ala Gly Val Ala Gly Leu Thr Thr
25       1               5                   10                  15
         gct ctt caa ctt ctt cgt aaa gga cat gag gtt aca att gtg tcc gag      96
         Ala Leu Gln Leu Leu Arg Lys Gly His Glu Val Thr Ile Val Ser Glu
                         20              25                  30
         ttt acg ccc ggt gat ctt agt atc gga tat acc tcg cct tgg gca ggt     144
30       Phe Thr Pro Gly Asp Leu Ser Ile Gly Tyr Thr Ser Pro Trp Ala Gly
                     35                  40                  45
         gcc aac tgg ctc aca ttt tac gat gga ggc aag tta gcc gac tac gat     192
         Ala Asn Trp Leu Thr Phe Tyr Asp Gly Gly Lys Leu Ala Asp Tyr Asp
                 50                  55                  60
35       gcc gtc tct tat cct atc ttg cga gag ctg gct cga agc agc ccc gag     240
         Ala Val Ser Tyr Pro Ile Leu Arg Glu Leu Ala Arg Ser Ser Pro Glu
         65                  70                  75                  80
         gct gga att cga ctc atc agc caa cgc tcc cat gtt ctc aag cgt gat     288
         Ala Gly Ile Arg Leu Ile Ser Gln Arg Ser His Val Leu Lys Arg Asp
40                           85                  90                  95
         ctt cct aaa ctg gaa gtt gcc atg tcg gcc atc tgt caa cgc aat ccc     336
         Leu Pro Lys Leu Glu Val Ala Met Ser Ala Ile Cys Gln Arg Asn Pro
                         100                 105                 110
         tgg ttc aaa aac aca gtc gat tct ttc gag att atc gag gac agg tcc     384
45       Trp Phe Lys Asn Thr Val Asp Ser Phe Glu Ile Ile Glu Asp Arg Ser
                     115                 120                 125
         agg att gtc cac gat gat gtg gct tat cta gtc gaa ttt cgt tcc gtt     432
```

```
      Arg Ile Val His Asp Asp Val Ala Tyr Leu Val Glu Phe Arg Ser Val
          130                     135                 140
      tgt atc cac acc gga gtc tac ttg aac tgg ctg atg tcc caa tgc tta      480
      Cys Ile His Thr Gly Val Tyr Leu Asn Trp Leu Met Ser Gln Cys Leu
 5    145                 150                     155                 160
      tcg ctc ggc gcc acg gtg gtt aaa cgt cga gtg aac cat atc aag gat      528
      Ser Leu Gly Ala Thr Val Val Lys Arg Arg Val Asn His Ile Lys Asp
                      165                 170                 175
      gcc aat tta cta cac tcc tca gga tca cgc ccc gac gtg att gtc aac      576
10    Ala Asn Leu Leu His Ser Ser Gly Ser Arg Pro Asp Val Ile Val Asn
                          180                 185                 190
      tgt agt ggt ctc ttt gcc cgg ttc ttg gga ggc gtc gag gac aag aag      624
      Cys Ser Gly Leu Phe Ala Arg Phe Leu Gly Gly Val Glu Asp Lys Lys
                      195                 200                 205
15    atg tac cct att cga gga caa gtc gtc ctt gtt cga aac tct ctt cct      672
      Met Tyr Pro Ile Arg Gly Gln Val Val Leu Val Arg Asn Ser Leu Pro
                  210                 215                 220
      ttt atg gcc tcc ttt tcc agc act cct gaa aaa gaa aat gaa gac gaa      720
      Phe Met Ala Ser Phe Ser Ser Thr Pro Glu Lys Glu Asn Glu Asp Glu
20    225                 230                 235                 240
      gct cta tat atc atg acc cga ttc gat ggt act tct atc att ggc ggt      768
      Ala Leu Tyr Ile Met Thr Arg Phe Asp Gly Thr Ser Ile Ile Gly Gly
                          245                 250                 255
      tgt ttc caa ccc aac aac tgg tca tcc gaa ccc gat cct tct ctc acc      816
25    Cys Phe Gln Pro Asn Asn Trp Ser Ser Glu Pro Asp Pro Ser Leu Thr
                      260                 265                 270
      cat cga atc ctg tct aga gcc ctc gac cga ttc ccg gaa ctg acc aaa      864
      His Arg Ile Leu Ser Arg Ala Leu Asp Arg Phe Pro Glu Leu Thr Lys
                      275                 280                 285
30    gat ggc cct ctt gac att gtg cgc gaa tgc gtt ggc cac cgt cct ggt      912
      Asp Gly Pro Leu Asp Ile Val Arg Glu Cys Val Gly His Arg Pro Gly
                  290                 295                     300
      aga gag ggc ggt ccc cga gta gaa tta gag aag atc ccc ggc gtt ggc      960
      Arg Glu Gly Gly Pro Arg Val Glu Leu Glu Lys Ile Pro Gly Val Gly
35    305                 310                     315                 320
      ttt gtt gtc cat aac tat ggt gcc gcc ggt gct ggt tac caa tcc tct     1008
      Phe Val Val His Asn Tyr Gly Ala Ala Gly Ala Gly Tyr Gln Ser Ser
                      325                     330                 335
      tac ggc atg gct gat gaa gct gtt tct tac gtc gaa aga gct ctt act     1056
40    Tyr Gly Met Ala Asp Glu Ala Val Ser Tyr Val Glu Arg Ala Leu Thr
                      340                 345                 350
      cgt cca aac ctt tag                                               1071
      Arg Pro Asn Leu
                  355

45
      <210>   8
      <211>   356
```

```
<212>  PRT
<213>  Trigonopsis variabilis


<400>  8

Met Ala Lys Ile Val Val Ile Gly Ala Gly Val Ala Gly Leu Thr Thr
1               5                   10                  15

Ala Leu Gln Leu Leu Arg Lys Gly His Glu Val Thr Ile Val Ser Glu
            20                  25                  30

Phe Thr Pro Gly Asp Leu Ser Ile Gly Tyr Thr Ser Pro Trp Ala Gly
            35                  40                  45

Ala Asn Trp Leu Thr Phe Tyr Asp Gly Gly Lys Leu Ala Asp Tyr Asp
        50                  55                  60

Ala Val Ser Tyr Pro Ile Leu Arg Glu Leu Ala Arg Ser Ser Pro Glu
65                  70                  75                  80

Ala Gly Ile Arg Leu Ile Ser Gln Arg Ser His Val Leu Lys Arg Asp
                85                  90                  95

Leu Pro Lys Leu Glu Val Ala Met Ser Ala Ile Cys Gln Arg Asn Pro
            100                 105                 110

Trp Phe Lys Asn Thr Val Asp Ser Phe Glu Ile Ile Glu Asp Arg Ser
            115                 120                 125

Arg Ile Val His Asp Asp Val Ala Tyr Leu Val Glu Phe Arg Ser Val
        130                 135                 140

Cys Ile His Thr Gly Val Tyr Leu Asn Trp Leu Met Ser Gln Cys Leu
145                 150                 155                 160

Ser Leu Gly Ala Thr Val Val Lys Arg Arg Val Asn His Ile Lys Asp
                165                 170                 175

Ala Asn Leu Leu His Ser Ser Gly Ser Arg Pro Asp Val Ile Val Asn
            180                 185                 190

Cys Ser Gly Leu Phe Ala Arg Phe Leu Gly Gly Val Glu Asp Lys Lys
            195                 200                 205

Met Tyr Pro Ile Arg Gly Gln Val Val Leu Val Arg Asn Ser Leu Pro
    210                 215                 220

Phe Met Ala Ser Phe Ser Ser Thr Pro Glu Lys Glu Asn Glu Asp Glu
```

225   230   235   240

Ala Leu Tyr Ile Met Thr Arg Phe Asp Gly Thr Ser Ile Ile Gly Gly
                    245   250   255

Cys Phe Gln Pro Asn Asn Trp Ser Ser Glu Pro Asp Pro Ser Leu Thr
                    260   265   270

His Arg Ile Leu Ser Arg Ala Leu Asp Arg Phe Pro Glu Leu Thr Lys
            275   280   285

Asp Gly Pro Leu Asp Ile Val Arg Glu Cys Val Gly His Arg Pro Gly
        290   295   300

Arg Glu Gly Gly Pro Arg Val Glu Leu Glu Lys Ile Pro Gly Val Gly
305   310   315   320

Phe Val Val His Asn Tyr Gly Ala Ala Gly Ala Gly Tyr Gln Ser Ser
                325   330   335

Tyr Gly Met Ala Asp Glu Ala Val Ser Tyr Val Glu Arg Ala Leu Thr
                340   345   350

Arg Pro Asn Leu
            355

<210> 9
<211> 1047
<212> DNA
<213> Schizosaccharomyces pombe

<220>
<221> CDS
<222> (22)..(1041)
<223> coding for DAAO

<400> 9
atgactaagg aaaataagcc a aga gat att gtc atc gtt ggc gct ggc gtt   51
                        Arg Asp Ile Val Ile Val Gly Ala Gly Val
                        1   5   10
att gga ttg acc act gct tgg att ctt tca gac ttg ggt ctt gct cct   99
Ile Gly Leu Thr Thr Ala Trp Ile Leu Ser Asp Leu Gly Leu Ala Pro
                15   20   25
cgt att aag gtg att gcc aag tat acg cct gaa gat cgt tct gta gaa   147
Arg Ile Lys Val Ile Ala Lys Tyr Thr Pro Glu Asp Arg Ser Val Glu
            30   35   40
tac act tcc cct tgg gct ggc gca aat ttc tgt agc att tct gct act   195

```
        Tyr Thr Ser Pro Trp Ala Gly Ala Asn Phe Cys Ser Ile Ser Ala Thr
                45                  50                  55
        gat gac aat gct ttg cgc tgg gat aaa atc act tac cat cgt ttc gcc     243
        Asp Asp Asn Ala Leu Arg Trp Asp Lys Ile Thr Tyr His Arg Phe Ala
  5         60                  65                  70
        tac ttg gcg aaa act cgt cct gaa gca gga atc cgt ttt gct gat ctt     291
        Tyr Leu Ala Lys Thr Arg Pro Glu Ala Gly Ile Arg Phe Ala Asp Leu
        75                  80                  85                  90
        cga gaa ttg tgg gag tac gag ccg aaa cac gac aaa atc aga tcc tgg     339
 10     Arg Glu Leu Trp Glu Tyr Glu Pro Lys His Asp Lys Ile Arg Ser Trp
                        95                  100                 105
        aat acc tat gtc aga gat ttc aaa gtt atc cct gaa aaa gat ctt cca     387
        Asn Thr Tyr Val Arg Asp Phe Lys Val Ile Pro Glu Lys Asp Leu Pro
                        110                 115                 120
 15     gga gaa tgt atc tac gga cat aag gcc acc acc ttt tta atc aac gct     435
        Gly Glu Cys Ile Tyr Gly His Lys Ala Thr Thr Phe Leu Ile Asn Ala
                        125                 130                 135
        cct cat tac ttg aat tat atg tac aag ctg ctc att gaa gct ggc gtc     483
        Pro His Tyr Leu Asn Tyr Met Tyr Lys Leu Leu Ile Glu Ala Gly Val
 20         140                 145                 150
        gaa ttt gaa aag aaa gaa ttg agt cac atc aaa gag act gtc gaa gaa     531
        Glu Phe Glu Lys Lys Glu Leu Ser His Ile Lys Glu Thr Val Glu Glu
        155                 160                 165                 170
        act cca gaa gct tca gta gta ttt aat tgc act ggt ctc tgg gct tcc     579
 25     Thr Pro Glu Ala Ser Val Val Phe Asn Cys Thr Gly Leu Trp Ala Ser
                        175                 180                 185
        aaa ttg ggt ggc gtt gaa gac ccg gac gtt tat ccg act cgt gga cat     627
        Lys Leu Gly Gly Val Glu Asp Pro Asp Val Tyr Pro Thr Arg Gly His
                        190                 195                 200
        gtt gtt ttg gtt aag gct cct cat gta aca gaa act cgc att ttg aat     675
 30     Val Val Leu Val Lys Ala Pro His Val Thr Glu Thr Arg Ile Leu Asn
                        205                 210                 215
        ggc aag aac tct gat acc tat att att cct cgt ccc tta aat ggt gga     723
        Gly Lys Asn Ser Asp Thr Tyr Ile Ile Pro Arg Pro Leu Asn Gly Gly
                220                 225                 230
        gtc att tgc ggc ggt ttc atg caa cca gga aac tgg gat cgt gaa att     771
        Val Ile Cys Gly Gly Phe Met Gln Pro Gly Asn Trp Asp Arg Glu Ile
        235                 240                 245                 250
        cac cct gaa gac act ttg gat atc ctt aag aga aca tcg gct ttg atg     819
 40     His Pro Glu Asp Thr Leu Asp Ile Leu Lys Arg Thr Ser Ala Leu Met
                        255                 260                 265
        cca gaa ttg ttc cac ggc aag ggt ccg gag ggt gct gaa att att caa     867
        Pro Glu Leu Phe His Gly Lys Gly Pro Glu Gly Ala Glu Ile Ile Gln
                        270                 275                 280
 45     gaa tgt gtc gga ttc cgt cct tct cga aag ggt ggt gcc cgc gta gag     915
        Glu Cys Val Gly Phe Arg Pro Ser Arg Lys Gly Gly Ala Arg Val Glu
                285                 290                 295
```

```
ctt gat gtt gtt ccc ggc acc tca gtc ccc ctt gtt cat gat tac ggt        963
Leu Asp Val Val Pro Gly Thr Ser Val Pro Leu Val His Asp Tyr Gly
    300                 305                 310

gct tct ggc aca gga tac caa gct ggt tat ggt atg gct ctt gac tct        1011
Ala Ser Gly Thr Gly Tyr Gln Ala Gly Tyr Gly Met Ala Leu Asp Ser
    315                 320                 325                 330

gtc atg ttg gct ctt cct aaa atc aaa ttg gcttag                          1047
Val Met Leu Ala Leu Pro Lys Ile Lys Leu
                    335                 340
```

<210> 10
<211> 340
<212> PRT
<213> Schizosaccharomyces pombe

<400> 10

```
Arg Asp Ile Val Ile Val Gly Ala Gly Val Ile Gly Leu Thr Thr Ala
1               5                   10                  15

Trp Ile Leu Ser Asp Leu Gly Leu Ala Pro Arg Ile Lys Val Ile Ala
            20                  25                  30

Lys Tyr Thr Pro Glu Asp Arg Ser Val Glu Tyr Thr Ser Pro Trp Ala
            35                  40                  45

Gly Ala Asn Phe Cys Ser Ile Ser Ala Thr Asp Asp Asn Ala Leu Arg
        50                  55                  60

Trp Asp Lys Ile Thr Tyr His Arg Phe Ala Tyr Leu Ala Lys Thr Arg
65                  70                  75                  80

Pro Glu Ala Gly Ile Arg Phe Ala Asp Leu Arg Glu Leu Trp Glu Tyr
                85                  90                  95

Glu Pro Lys His Asp Lys Ile Arg Ser Trp Asn Thr Tyr Val Arg Asp
            100                 105                 110

Phe Lys Val Ile Pro Glu Lys Asp Leu Pro Gly Glu Cys Ile Tyr Gly
            115                 120                 125

His Lys Ala Thr Thr Phe Leu Ile Asn Ala Pro His Tyr Leu Asn Tyr
        130                 135                 140

Met Tyr Lys Leu Leu Ile Glu Ala Gly Val Glu Phe Glu Lys Lys Glu
145                 150                 155                 160

Leu Ser His Ile Lys Glu Thr Val Glu Glu Thr Pro Glu Ala Ser Val
```

                     165   170   175

      Val Phe Asn Cys Thr Gly Leu Trp Ala Ser Lys Leu Gly Gly Val Glu
                  180   185   190
  5

      Asp Pro Asp Val Tyr Pro Thr Arg Gly His Val Val Leu Val Lys Ala
                  195   200   205

      Pro His Val Thr Glu Thr Arg Ile Leu Asn Gly Lys Asn Ser Asp Thr
 10   210   215   220

      Tyr Ile Ile Pro Arg Pro Leu Asn Gly Gly Val Ile Cys Gly Gly Phe
      225   230   235   240

 15   Met Gln Pro Gly Asn Trp Asp Arg Glu Ile His Pro Glu Asp Thr Leu
                       245   250   255

      Asp Ile Leu Lys Arg Thr Ser Ala Leu Met Pro Glu Leu Phe His Gly
                  260   265   270
 20
      Lys Gly Pro Glu Gly Ala Glu Ile Ile Gln Glu Cys Val Gly Phe Arg
                  275   280   285

      Pro Ser Arg Lys Gly Gly Ala Arg Val Glu Leu Asp Val Val Pro Gly
 25   290   295   300

      Thr Ser Val Pro Leu Val His Asp Tyr Gly Ala Ser Gly Thr Gly Tyr
      305   310   315   320

 30   Gln Ala Gly Tyr Gly Met Ala Leu Asp Ser Val Met Leu Ala Leu Pro
                       325   330   335

      Lys Ile Lys Leu
                  340
 35
      <210>   11
      <211>   963
      <212>   DNA
      <213>   Streptomyces coelicolor
 40
      <220>
      <221>   CDS
      <222>   (31)..(957)
      <223>   coding for DAAO
 45
      <220>
      <221>   misc_feature

<222>   (880)..(936)
<223>   DAAO signature

<400>   11

gtggaaaccg aactggatga cgagcgggat ggc gaa gtc gtc gtg gtc ggc ggc   54
                                     Gly Glu Val Val Val Val Gly Gly
                                      1   5

ggg gtg atc ggg ctg acg acg gcc gtc gtc ctc gcc gag cgg ggc aga   102
Gly Val Ile Gly Leu Thr Thr Ala Val Val Leu Ala Glu Arg Gly Arg
     10   15   20

cgg gtg cgg ctg tgg acc cgg gag ccc gcg gag cgg acc acc tcg gtg   150
Arg Val Arg Leu Trp Thr Arg Glu Pro Ala Glu Arg Thr Thr Ser Val
 25   30   -35   40

gta gcg ggc ggg ctg tgg tgg ccg tac cgg atc gag ccg gtc gcg ctg   198
Val Ala Gly Gly Leu Trp Trp Pro Tyr Arg Ile Glu Pro Val Ala Leu
             45   50   55

gcc cag gcc tgg gcg ctg cgt tcc ctg gac gtg tac gag gag ctg gcg   246
Ala Gln Ala Trp Ala Leu Arg Ser Leu Asp Val Tyr Glu Glu Leu Ala
             60   65   70

gca cgg ccc ggg cag acc ggc gta cgc atg ctc gaa ggg gtg ctc ggc   294
Ala Arg Pro Gly Gln Thr Gly Val Arg Met Leu Glu Gly Val Leu Gly
             75   80   85

gag acc ggc ctg gac gag gtg gac ggg tgg gcc gcg gcc cgg ctg ccg   342
Glu Thr Gly Leu Asp Glu Val Asp Gly Trp Ala Ala Ala Arg Leu Pro
     90   95   100

ggg ctg cgc gcg gcg agc gcc gcc gag tac gcc ggg acg ggg ctg tgg   390
Gly Leu Arg Ala Ala Ser Ala Ala Glu Tyr Ala Gly Thr Gly Leu Trp
105   110   115   120

gcg cgg ctg ccg ctc atc gac atg tcg acc cat ctg ccg tgg ctg cgg   438
Ala Arg Leu Pro Leu Ile Asp Met Ser Thr His Leu Pro Trp Leu Arg
                 125   130   135

gag cgg ctg ctg gcc gcg ggc ggc acg gtg gag gac cgc gcg gtg acc   486
Glu Arg Leu Leu Ala Ala Gly Gly Thr Val Glu Asp Arg Ala Val Thr
                 140   145   150

gat ctg gcc gag gcg gac gcg ccg gtg gtg gtc aac tgc acc ggc ctg   534
Asp Leu Ala Glu Ala Asp Ala Pro Val Val Val Asn Cys Thr Gly Leu
         155   160   165

ggc gcc cgg gag ctg gtg ccg gac ccg gcg gta cgg ccg gtg cgc gga   582
Gly Ala Arg Glu Leu Val Pro Asp Pro Ala Val Arg Pro Val Arg Gly
         170   175   180

cag ctg gtc gtc gtg gag aac ccc ggc atc cac aac tgg ctg gtc gcg   630
Gln Leu Val Val Val Glu Asn Pro Gly Ile His Asn Trp Leu Val Ala
185   190   195   200

gcc gac gcg gac tcc ggg gag acg acg tac ttc ctt ccg cag ccg gga   678
Ala Asp Ala Asp Ser Gly Glu Thr Thr Tyr Phe Leu Pro Gln Pro Gly
                 205   210   215

cgg ctc ctg ctg ggc ggc acg gct gag gag gac gcc tgg tcg acc gag   726

```
         Arg Leu Leu Leu Gly Gly Thr Ala Glu Glu Asp Ala Trp Ser Thr Glu
                     220             225             230
         ccg gac ccg gag gtc gcg gcg gcc atc gtg cga cgg tgc gcg gcc ctg        774
         Pro Asp Pro Glu Val Ala Ala Ala Ile Val Arg Arg Cys Ala Ala Leu
  5              235             240             245
         cgt ccc gag atc gcc gga gcg cgg gtg ctc gcg cac ctg gtg ggg ctg        822
         Arg Pro Glu Ile Ala Gly Ala Arg Val Leu Ala His Leu Val Gly Leu
                 250             255             260
         cgg ccg gcc cgg gac gcg gtc cgg ctg gag cgc ggg acg ctg ccg gac        870
 10      Arg Pro Ala Arg Asp Ala Val Arg Leu Glu Arg Gly Thr Leu Pro Asp
         265             270             275             280
         ggg cgc cgg ctg gtg cac aac tac ggt cac ggc ggc gcg ggc gtc acc        918
         Gly Arg Arg Leu Val His Asn Tyr Gly His Gly Gly Ala Gly Val Thr
                     285             290             295
 15      gtg gcc tgg ggc tgc gct cag gag gcg gcc cgg ctc gcc tcctga           963
         Val Ala Trp Gly Cys Ala Gln Glu Ala Ala Arg Leu Ala
                     300             305
```

         <210>   12
 20   <211>   309
         <212>   PRT
         <213>   Streptomyces coelicolor

         <220>
 25   <221>   misc_feature
         <222>   (880)..(936)
         <223>   DAAO signature

         <400>   12
 30   Gly Glu Val Val Val Val Gly Gly Gly Val Ile Gly Leu Thr Thr Ala
         1   5   10   15

         Val Val Leu Ala Glu Arg Gly Arg Arg Val Arg Leu Trp Thr Arg Glu
                     20   25   30
 35
         Pro Ala Glu Arg Thr Thr Ser Val Val Ala Gly Gly Leu Trp Trp Pro
                     35   40   45

         Tyr Arg Ile Glu Pro Val Ala Leu Ala Gln Ala Trp Ala Leu Arg Ser
 40   50   55   60

         Leu Asp Val Tyr Glu Glu Leu Ala Ala Arg Pro Gly Gln Thr Gly Val
         65   70   75   80

 45   Arg Met Leu Glu Gly Val Leu Gly Glu Thr Gly Leu Asp Glu Val Asp
                     85   90   95

```
       Gly Trp Ala Ala Ala Arg Leu Pro Gly Leu Arg Ala Ala Ser Ala Ala
                   100                 105                 110


       Glu Tyr Ala Gly Thr Gly Leu Trp Ala Arg Leu Pro Leu Ile Asp Met
  5                115                 120                 125


       Ser Thr His Leu Pro Trp Leu Arg Glu Arg Leu Leu Ala Ala Gly Gly
           130                 135                 140


 10    Thr Val Glu Asp Arg Ala Val Thr Asp Leu Ala Glu Ala Asp Ala Pro
       145                 150                 155                 160


       Val Val Val Asn Cys Thr Gly Leu Gly Ala Arg Glu Leu Val Pro Asp
                   165                 170                 175
 15

       Pro Ala Val Arg Pro Val Arg Gly Gln Leu Val Val Val Glu Asn Pro
                   180                 185                 190


       Gly Ile His Asn Trp Leu Val Ala Ala Asp Ala Asp Ser Gly Glu Thr
 20                195                 200                 205


       Thr Tyr Phe Leu Pro Gln Pro Gly Arg Leu Leu Leu Gly Gly Thr Ala
           210                 215                 220


 25    Glu Glu Asp Ala Trp Ser Thr Glu Pro Asp Pro Glu Val Ala Ala Ala
       225                 230                 235                 240


       Ile Val Arg Arg Cys Ala Ala Leu Arg Pro Glu Ile Ala Gly Ala Arg
                   245                 250                 255
 30

       Val Leu Ala His Leu Val Gly Leu Arg Pro Ala Arg Asp Ala Val Arg
                   260                 265                 270


       Leu Glu Arg Gly Thr Leu Pro Asp Gly Arg Arg Leu Val His Asn Tyr
 35                275                 280                 285


       Gly His Gly Gly Ala Gly Val Thr Val Ala Trp Gly Cys Ala Gln Glu
           290                 295                 300


 40    Ala Ala Arg Leu Ala
       305


       <210>   13
       <211>   1038
 45    <212>   DNA
       <213>   Candida boidinii
```

<220>
<221> CDS
<222> (1)..(1035)
<223> coding for DAAO

<400> 13

```
atg ggt gat caa att gtt gtt ctt ggt tcc ggt att att ggt tta tat     48
Met Gly Asp Gln Ile Val Val Leu Gly Ser Gly Ile Ile Gly Leu Tyr
1               5                   10                  15
act aca tac tgt tta atc tat gag gct gga tgt gct cca gct aaa att     96
Thr Thr Tyr Cys Leu Ile Tyr Glu Ala Gly Cys Ala Pro Ala Lys Ile
                20                  25                  30
act att gtt gct gaa ttt tta cca ggt gat caa tct aca tta tat aca    144
Thr Ile Val Ala Glu Phe Leu Pro Gly Asp Gln Ser Thr Leu Tyr Thr
            35                  40                  45
tct cca tgg gca ggt ggt aat ttt tct tgt att tca cca gct gat gat    192
Ser Pro Trp Ala Gly Gly Asn Phe Ser Cys Ile Ser Pro Ala Asp Asp
        50                  55                  60
aca aca ttg gct tat gat aaa ttc aca tat ctt aat tta ttc aag att    240
Thr Thr Leu Ala Tyr Asp Lys Phe Thr Tyr Leu Asn Leu Phe Lys Ile
65                  70                  75                  80
cac aaa aaa tta ggt gga cca gaa tgt gga tta gat aat aag cca agt    288
His Lys Lys Leu Gly Gly Pro Glu Cys Gly Leu Asp Asn Lys Pro Ser
                85                  90                  95
act gaa tat tgg gat ttt tat cct ggt gat gaa aaa gtc aat tct tta    336
Thr Glu Tyr Trp Asp Phe Tyr Pro Gly Asp Glu Lys Val Asn Ser Leu
                100                 105                 110
aaa caa tat ctt aaa gat ttt aaa gtt att cca aaa tca gaa tta cca    384
Lys Gln Tyr Leu Lys Asp Phe Lys Val Ile Pro Lys Ser Glu Leu Pro
            115                 120                 125
gaa ggt gtt gaa tat ggt att agt tat act aca tgg aat ttc aac tgt    432
Glu Gly Val Glu Tyr Gly Ile Ser Tyr Thr Thr Trp Asn Phe Asn Cys
        130                 135                 140
cct gtt ttc tta caa aat atg gct aat ttt tta aat aaa aga aat gtt    480
Pro Val Phe Leu Gln Asn Met Ala Asn Phe Leu Asn Lys Arg Asn Val
145                 150                 155                 160
acc att att aga aaa cat tta aca cat att tct caa gct tat tta aca    528
Thr Ile Ile Arg Lys His Leu Thr His Ile Ser Gln Ala Tyr Leu Thr
                165                 170                 175
gtt aat aca aaa gtt gtt ttc aac tgt aca ggt att ggt gct gct gat    576
Val Asn Thr Lys Val Val Phe Asn Cys Thr Gly Ile Gly Ala Ala Asp
                180                 185                 190
tta ggt ggt gtt aaa gat gaa aaa gtt tat cca act aga gga caa gtt    624
Leu Gly Gly Val Lys Asp Glu Lys Val Tyr Pro Thr Arg Gly Gln Val
            195                 200                 205
gtt gtt gtt aga gct cca cat att caa gaa aat aaa atg aga tgg ggt    672
Val Val Val Arg Ala Pro His Ile Gln Glu Asn Lys Met Arg Trp Gly
```

```
        210                   215                   220
aaa gac tat gct act tat att att cca aga cca tat tct aat ggt gaa        720
Lys Asp Tyr Ala Thr Tyr Ile Ile Pro Arg Pro Tyr Ser Asn Gly Glu
225                   230                   235                   240
tta gtc tta ggt ggt ttc tta caa aag gat aat tgg aca ggt aat act       768
Leu Val Leu Gly Gly Phe Leu Gln Lys Asp Asn Trp Thr Gly Asn Thr
                    245                   250                   255
ttt ggt ttt gaa act gat gat att gtt agt aga act aca tct tta tta       816
Phe Gly Phe Glu Thr Asp Asp Ile Val Ser Arg Thr Thr Ser Leu Leu
                260                   265                   270
cca aag att tta gat gaa cca ctt cat att att aga gtt gca gct ggt       864
Pro Lys Ile Leu Asp Glu Pro Leu His Ile Ile Arg Val Ala Ala Gly
            275                   280                   285
tta aga cca agt aga cat ggt ggt cca aga att gaa gct gaa gtt tgt       912
Leu Arg Pro Ser Arg His Gly Gly Pro Arg Ile Glu Ala Glu Val Cys
        290                   295                   300
gaa gaa ggt aaa tta act att cat aat tat ggt gct tct gga tat ggt       960
Glu Glu Gly Lys Leu Thr Ile His Asn Tyr Gly Ala Ser Gly Tyr Gly
305                   310                   315                   320
tat caa gct ggt tat ggt atg tct tat gaa gct gtc aaa ctt tta gtt      1008
Tyr Gln Ala Gly Tyr Gly Met Ser Tyr Glu Ala Val Lys Leu Leu Val
                    325                   330                   335
gat aac caa aaa gtt aaa gct aaa ctt tag                              1038
Asp Asn Gln Lys Val Lys Ala Lys Leu
                340                   345


<210>  14
<211>  345
<212>  PRT
<213>  Candida boidinii


<400>  14
Met Gly Asp Gln Ile Val Val Leu Gly Ser Gly Ile Ile Gly Leu Tyr
1               5                   10                  15


Thr Thr Tyr Cys Leu Ile Tyr Glu Ala Gly Cys Ala Pro Ala Lys Ile
            20                  25                  30


Thr Ile Val Ala Glu Phe Leu Pro Gly Asp Gln Ser Thr Leu Tyr Thr
            35                  40                  45


Ser Pro Trp Ala Gly Gly Asn Phe Ser Cys Ile Ser Pro Ala Asp Asp
    50                  55                  60


Thr Thr Leu Ala Tyr Asp Lys Phe Thr Tyr Leu Asn Leu Phe Lys Ile
65                  70                  75                  80
```

                His Lys Lys Leu Gly Gly Pro Glu Cys Gly Leu Asp Asn Lys Pro Ser
                                85   90   95

         5   Thr Glu Tyr Trp Asp Phe Tyr Pro Gly Asp Glu Lys Val Asn Ser Leu
                            100   105   110

                Lys Gln Tyr Leu Lys Asp Phe Lys Val Ile Pro Lys Ser Glu Leu Pro
                            115   120   125

        10   Glu Gly Val Glu Tyr Gly Ile Ser Tyr Thr Thr Trp Asn Phe Asn Cys
                        130   135   140

                Pro Val Phe Leu Gln Asn Met Ala Asn Phe Leu Asn Lys Arg Asn Val
                145   150   155   160
        15

                Thr Ile Ile Arg Lys His Leu Thr His Ile Ser Gln Ala Tyr Leu Thr
                                165   170   175

        20   Val Asn Thr Lys Val Val Phe Asn Cys Thr Gly Ile Gly Ala Ala Asp
                                180   185   190

                Leu Gly Gly Val Lys Asp Glu Lys Val Tyr Pro Thr Arg Gly Gln Val
                        195   200   205

        25   Val Val Val Arg Ala Pro His Ile Gln Glu Asn Lys Met Arg Trp Gly
                        210   215   220

                Lys Asp Tyr Ala Thr Tyr Ile Ile Pro Arg Pro Tyr Ser Asn Gly Glu
                225   230   235   240
        30

                Leu Val Leu Gly Gly Phe Leu Gln Lys Asp Asn Trp Thr Gly Asn Thr
                                245   250   255

        35   Phe Gly Phe Glu Thr Asp Asp Ile Val Ser Arg Thr Thr Ser Leu Leu
                            260   265   270

                Pro Lys Ile Leu Asp Glu Pro Leu His Ile Ile Arg Val Ala Ala Gly
                            275   280   285

        40   Leu Arg Pro Ser Arg His Gly Gly Pro Arg Ile Glu Ala Glu Val Cys
                        290   295   300

                Glu Glu Gly Lys Leu Thr Ile His Asn Tyr Gly Ala Ser Gly Tyr Gly
                305   310   315   320
        45

                Tyr Gln Ala Gly Tyr Gly Met Ser Tyr Glu Ala Val Lys Leu Leu Val
                                325   330   335

```
         Asp Asn Gln Lys Val Lys Ala Lys Leu
                     340                 345


  5      <210>  15
         <211>  12334
         <212>  DNA
         <213>  STPT GUS Nit-P daao


 10      <220>
         <221>  misc_feature
         <222>  (38)..(183)
         <223>  Agrobacterium Right Border


 15      <220>
         <221>  terminator
         <222>  (384)..(639)
         <223>  complementary: NOS terminator


 20      <220>
         <221>  misc_feature
         <222>  (716)..(1822)
         <223>  complementary: coding for Rhodotorula gracilis DAAO


 25      <220>
         <221>  promoter
         <222>  (1866)..(3677)
         <223>  complementary: A.thaliana nitrilase I promoter


 30      <220>
         <221>  terminator
         <222>  (3767)..(3971)
         <223>  complementary: 35S terminator


 35      <220>
         <221>  misc_feature
         <222>  (4046)..(6043)
         <223>  complementary: coding for beta-glucuronidase (comprising intron)


 40      <220>
         <221>  promoter
         <222>  (6097)..(7414)
         <223>  complementary: sTPT promoter


 45      <220>
         <221>  misc_feature
```

```
<222>    (7486)..(7702)
<223>    Agrobacterium Left Border


<400>    15
```

```
aatattcaaa caaacacata cagcgcgact tatcatggac atacaaatgg acgaacggat      60
aaacctttc acgcccttt aaatatccga ttattctaat aaacgctctt ttctcttagg       120
tttacccgcc aatatatcct gtcaaacact gatagtttaa actgaaggcg ggaaacgaca     180
atcagatctg gtaccttgcg ccgggtaccc caaactgtct cacgacgttt tgaacccaga     240
ttaccctgtt atccctagtc gagcggccgc cagtgtgatg gatatctgca gaattcgccc     300
ttttagatca gcacactggc ggccgttact agtggatcaa ttcactggcc gtcgttttac     360
aacgactcag agcttgacag gaggcccgat ctagtaacat agatgacacc gcgcgcgata     420
atttatccta gtttgcgcgc tatattttgt tttctatcgc gtattaaatg tataattgcg     480
ggactctaat cataaaaacc catctcataa ataacgtcat gcattacatg ttaattatta     540
catgcttaac gtaattcaac agaaattata tgataatcat cgcaagaccg gcaacaggat     600
tcaatcttaa gaaactttat tgccaaatgt ttgaacgatc gggatcatc cgggtctgtg      660
gcgggaactc cacgaaaata tccgaacgca gcaagatcta gagcttgggt cccgcctaca     720
acttcgactc ccgcgccgcg ccgtggtacc gctggaacgc ctcgtcgacg agctgcgcga     780
catcctccgc cgcgccccaa ctctgctggt atcccgcact cgagaagcca tacgcatgca     840
caagcgtgac ctccttctcc ttcgccgctc gtgcgctgcc cctgccgagc gagaggggcg     900
actttgtccg gtcgagaggc aggacgatcc gttctgcctc aacgcggggt ccgcctcgtc     960
gtgcaggtcg caagccgacg ttgtggcgga ggacctcgat gccttcgatc gttccgtcgc    1020
tcgagatggt cgggtcgagg cgcaagcagt gcttgaggat ccgctggacc gtctctgggt    1080
tgacagacaa gtcccagtct cccacgccgt acgtcccgcc gcagatgact tcgccacctg    1140
gtcggggaat gatgtaggcg ggagaagcgg ggtcggacga gtccatcgtg catcgcttgc    1200
atgggacttt gacgaggacg gtttgcccgc ggattggctc ggcggcttgg tcgtcgatgc    1260
ccgcaatcga cttggcgcca agtcccgtag cgttgaccac caaatccgca ccgtcgaacg    1320
cctgctcaag cgacgtaacg gtccgtctct caaacgtcgc gccgagcttc tgcagctctc    1380
ttgcaaggta ctggcagtac tttggtgcgt ggacggagag ggtgtcgtag ttacgccga     1440
tagcgccagg tggacattcg gaagatggga gggggcggta atttggcgtg atgtccttgt    1500
accagtgccc gagcaagccg tcttcgttct gcgcgaaccg cctcgtcccc ttgagccaca    1560
tggcatggcc cgtcgggacc aactcgaccc acttcttgaa agtcgattct tcccatttg     1620
cttgtcgagg accgtctgta agcgtcatga aaggcgtcca attcgcgcca gcccatggtg    1680
aagcgaaagt ctggctcgag acgtcctccg gcaagtcgcg cgcgagaata tgcacgctgt    1740
agcccttccg agcgaggatg agggcgctgc tcagaccgat aacgcctgat ccgaggacaa    1800
cgacgcgctt ctgcgagtgc atgggccctc gactagagtc gagatccgat atcgcccggg    1860
ctcgagtctt tgttttttac tttggttcat gacactcaga gacttgagag aagcaatata    1920
tagacttttt tttgttttt ttttgtggtc acgtttattt tcctattgga gacggtaacg     1980
aagatcgaac ctgtggtgga aatgaaacaa ggtgggacta gcccacgtgg tttcttttct    2040
ctgcattgat ttgttttgt tttttttgta aagttcacat caaacctact aataattgag     2100
aagaaaata aaatctattg attgattaaa ccagccgatg ctttatgtct gaatataaaa     2160
aagaagtgaa aaccccgttt aagaattaca acggtggttt acaagtatt tggacacaat      2220
aaatccaaac gaaataaaac aaaatggaga actaccaaat aaaaaacaaa taaaaaactt     2280
aaaagaattt attccatttt ttttcccgta gaatttattc ttttatggat ccttaaatc     2340
catatttgat gcattttgat cctcataat aggtaataat atatactatg ttatagatat      2400
gtttctaatt cgtattaacc tacctttttt tggtcgtacg attctaccta ataatattga    2460
acggaattga tgtttggac cacttagaaa gtatttttt tttggtttgt cttagctgta       2520
tttcattaaa tataaattta aataagaaat gtcataaata aaatttgacg tatagatttt    2580
```

```
ttaaatccat tttatgttat ttaatatttg aaatgtgagt ttggctccta tttaatctta   2640
ggatgggtta atactaagtt ttccttaatg aattatctca gagaaactgg attaaataaa   2700
ctaaaaaata gatcaatgtg ttttggtccg gtcaaatatc tttggattta ctattattgg   2760
cgaaaagaaa gtctcatata gtaaatcata ttcctacaag agaaatcaaa attttgaat   2820
taacatggat tgtatagttt cttatataac caattagttc gcatcaagaa aaccaaaccc   2880
caattaataa tcaaacgggc ttggtaggaa tatttcattg cagctttcag ataaaagaaa   2940
aaaacacaca ctcaagtctt ttatttcatc tttcttactt gcaggaactc aaattccact   3000
ttgccacttt tctttacaaa taaacacaaa ttgtcaatga aacgaaatag tctttttatg   3060
caaacactgt ttgtcttttt tcgatcacgt ttctgattgt gacagccatc catatatata   3120
gggaatgtaa aacaacaaca tgtgaagtca catatacgta atggtttagc atagcttcta   3180
ttttcgttgt caatattagt cattccaaaa catttttaag aaaaataaat taatatatgt   3240
atattcttgg aactaatgta tgtggaaata cagtaactta attattaaac attctaaatg   3300
caaatatgca aagaaaaaaa agaaaagaac acaactgaaa tcaaagccag attcataata   3360
attggctaca tggttgtaga atgtagggta acacaacatc cagaattgaa cactcaaatt   3420
ggatgataga tggataatct ttagatacaa gagaattggt tctcttccat tattaacgaa   3480
aataaagaaa aaaagtttag cataaaagtt tgaaactcaa cataacattt tgaacttgac   3540
tccttcatag gagtgacatg aactgacgaa tcacaaccga ttacttgttt gagtcatctt   3600
ccgctttctc caccttcgaa atgaatgtga ccggtttctt cgggtgctca tttacggtca   3660
agtgtaaaac atctggtctc gaggtacctg gtagggataa cagggtaatc tgggttcaaa   3720
acgtcgtgag acagtttggt gcaggtcgaa attcgagctc ggtacccggt cactggattt   3780
tggttttagg aattagaaat tttattgata gaagtatttt acaaatacaa atacatacta   3840
agggtttctt atatgctcaa cacatgagcg aaaccctata agaaccctaa ttcccttatc   3900
tgggaactac tcacacatta ttctggagaa aaatagagag agatagattt gtagagagag   3960
actggtgatt tttgcgccgg gtaccgagct cggtagcaat tcccgaggct gtagccgacg   4020
atggtgcgcc aggagagttg ttgattcatt gtttgcctcc ctgctgcggt ttttcaccga   4080
agttcatgcc agtccagcgt ttttgcagca gaaaagccgc cgacttcggt ttgcggtcgc   4140
gagtgaagat ccctttcttg ttaccgccaa cgcgcaatat gccttgcgag gtcgcaaaat   4200
cggcgaaatt ccatacctgt tcaccgacga cggcgctgac gcgatcaaag acgcggtgat   4260
acatatccag ccatgcacac tgatactctt cactccacat gtcggtgtac attgagtgca   4320
gcccggctaa cgtatccacg ccgtattcgg tgatgataat cggctgatgc agtttctcct   4380
gccaggccag aagttctttt tccagtacct tctctgccgt ttccaaatcg ccgctttgga   4440
cataccatcc gtaataacgg ttcaggcaca gcacatcaaa gagatcgctg atggtatcgg   4500
tgtgagcgtc gcagaacatt acattgacgc aggtgatcgg acgcgtcggg tcgagtttac   4560
gcgttgcttc cgccagtggc gaaatattcc cgtgcacttg cggacgggta tccggttcgt   4620
tggcaatact ccacatcacc acgcttgggt ggttttgtc acgcgctatc agctctttaa   4680
tcgcctgtaa gtgcgcttgc tgagtttccc cgttgactgc ctcttcgctg tacagttctt   4740
tcggcttgtt gcccgcttcg aaaccaatgc ctaaagagag gttaaagccg acagcagcag   4800
tttcatcaat caccacgatg ccatgttcat ctgcccagtc gagcatctct tcagcgtaag   4860
ggtaatgcga ggtacggtag gagttggccc caatccagtc cattaatgcg tggtcgtgca   4920
ccatcagcac gttatcgaat cctttgccac gtaagtccgc atcttcatga cgaccaaagc   4980
cagtaaagta gaacggtttg tggttaatca ggaactgttc gcccttcact gccactgacc   5040
ggatgccgac gcgaagcggg tagatatcac actctgtctg gcttttggct gtgacgcaca   5100
gttcatagag ataaccttca cccggttgcc agaggtgcgg attcaccact tgcaaagtcc   5160
cgctagtgcc ttgtccagtt gcaaccacct gttgatccgc atcacgcagt caacgctga    5220
catcaccatt ggccaccacc tgccagtcaa cagacgcgtg gttacagtct gcgcgacat    5280
gcgtcaccac ggtgatatcg tccacccagg tgttcggcgt ggtgtagagc attacgctgc   5340
gatggattcc ggcatagtta aagaaatcat ggaagtaaga ctgcttttc ttgccgtttt    5400
```

```
cgtcggtaat caccattccc ggcgggatag tctgccagtt cagttcgttg ttcacacaaa    5460
cggtgatacc tgcacatcaa caaattttgg tcatatatta gaaaagttat aaattaaaat    5520
atacacactt ataaactaca gaaaagcaat tgctatatac tacattcttt tattttgaaa    5580
aaaatatttg aaatattata ttactactaa ttaatgataa ttattatata tatatcaaag    5640
gtagaagcag aaacttacgt acacttttcc cggcaataac atacggcgtg acatcggctt    5700
caaatggcgt atagccgccc tgatgctcca tcacttcctg attattgacc cacactttgc    5760
cgtaatgagt gaccgcatcg aaacgcagca cgatacgctg gcctgcccaa cctttcggta    5820
taaagacttc gcgctgatac cagacgttgc ccgcataatt acgaatatct gcatcggcga    5880
actgatcgtt aaaactgcct ggcacagcaa ttgcccggct ttcttgtaac gcgctttccc    5940
accaacgctg accaattcca cagttttcgc gatccagact gaatgcccac aggccgtcga    6000
gtttttttgat ttcacgggtt ggggtttcta caggacgtac catggtcgat cgactctaga    6060
ctagtggatc cgatatcgcc cgggctcgac tctagatgaa atcgaaattc agagttttga    6120
tagtgagagc aaagagggac ggacttatga ggatttcgag tatttcaaga gatggtactt    6180
gttgatcgga cggctacgat gatctcgatt tggttaatcc agtatctcgc ggtgtatgga    6240
gttatggtag ggttaatggt caatttcatc taacggtaga gaatgatgta attagataag    6300
aatcttgaga tactggttta gattggatga gtgtagggtc catcttatct tgataagtgg    6360
atggttttta gagacacagt gaatattagc caatcgaagt tccatatcac catcatcatc    6420
tgtataattt tgttttttttg gaagataata atgattgaaa ttttggtaga ttttatttt    6480
cattatttac cttgtatgtt gagtggtctt caaattattg aacgtgacag attcacaaga    6540
aagtagattt tttataaatg aaattttact tatttttaaag gtatctctat ttaatttctt    6600
ttgtttatgg ttgtctgtca gcatttgact tgcagtttca tgctcatagt catatacgtt    6660
attctaggct tttttgaata tcttattact tttttcgtaa tacaatttta taatttttatc    6720
aaagttatac aactataact aaaattaggg ttttctacaa aacaaaaaaa tcttctaatt    6780
ttttttgttg tagccagttt actcgtaagt tacaaaaaaa tacaaatgaa cccacatgta    6840
ttatgcgttt aactaggatt accatgtact ttcatgtact caattcaccc tatactcttt    6900
ttttttttttt ttctagttcc acccaatcta taaaattctg tccatttgac caaattcaat    6960
taatttctgt aattgcgatt taaaattaat attacatgtt cactatttct cgatttgagg    7020
gaacccgagt ttaaatatga taaaaatgtt gacccatcac tacaaatatg ttatagttta    7080
tacttaatag tggtgttttt ggggataatt gatgaattaa gtaaacatga ttcttcttat    7140
gaagttgatt gagtgattat tgtatgtaaa cctatgtgat tgatgttatt ggttgattga    7200
gtgattattg tattagtatg taagcaaaga tgattgttct tatgaggtaa tttgttactc    7260
attcatcctt ttgcatatga gaaattgtgt tagcgtacgc aaaacaatag agaacataaa    7320
agatatgtgt atttatttaa ggtgactttt gttaatgata ttgtagtatc tatacatta    7380
tatataactt gttgaatttg agtataagct atcaggatcc ggggatcct ctagagtcga    7440
ggtaccgagc tcgaattcac tggccgtcgt tttacaacga ctcagtactg cttggtaata    7500
attgtcatta gattgttttt atgcatagat gcactcgaaa tcagccaatt ttagacaagt    7560
atcaaacgga tgttaattca gtacattaaa gacgtccgca atgtgttatt aagttgtcta    7620
agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca gctccccgac    7680
cggcagctcg gcacaaaatc accacgcgtc taaaaggtg atgtgtattt gagtaaaaca    7740
gcttgcgtca tgcggtcgct gcgtatatga tgcgatgagt aaataaacaa atacgcaagg    7800
ggaacgcatg aaggttatcg ctgtacttaa ccagaaaggc gggtcaggca agacgaccat    7860
cgcaacccat ctagcccgcg ccctgcaact cgccggggcc gatgttctgt tagtcgattc    7920
cgatccccag ggcagtgccc gcgattgggc ggccgtgcgg gaagatcaac cgctaaccgt    7980
tgtcggcatc gaccgcccga cgattgaccg cgacgtgaag gccatcggcc ggcgcgactt    8040
cgtagtgatc gacggagcgc cccaggcggc ggacttggct gtgtccgcga tcaaggcagc    8100
cgacttcgtg ctgattccgg tgcagccaag cccttacgac atatgggcca ccgccgacct    8160
ggtggagctg gttaagcagc gcattgaggt cacggatgga aggctacaag cggcctttgt    8220
```

```
     cgtgtcgcgg gcgatcaaag gcacgcgcat cggcggtgag gttgccgagg cgctggccgg   8280
     gtacgagctg cccattcttg agtcccgtat cacgcagcgc gtgagctacc caggcactgc   8340
     cgccgccggc acaaccgttc ttgaatcaga acccgagggc gacgctgccc gcgaggtcca   8400
     ggcgctggcc gctgaaatta aatcaaaact catttgagtt aatgaggtaa agagaaaatg   8460
 5   agcaaaagca caaacacgct aagtgccggc cgtccgagcg cacgcagcag caaggctgca   8520
     acgttggcca gcctggcaga cacgccagcc atgaagcggg tcaactttca gttgccggcg   8580
     gaggatcaca ccaagctgaa gatgtacgcg gtacgccaag gcaagaccat taccgagctg   8640
     ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa tgagtagatg   8700
     aattttagcg gctaaaggag gcggcatgga aaatcaagaa caaccaggca ccgacgccgt   8760
10   ggaatgcccc atgtgtggag gaacgggcgg ttggccaggc gtaagcggct gggttgtctg   8820
     ccggccctgc aatggcactg gaacccccaa gcccgaggaa tcggcgtgag cggtcgcaaa   8880
     ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag   8940
     gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag acgccccggt gaatcgtggc   9000
     aagggggccgc tgatcgaatc cgcaaagaat cccggcaacc gccggcagcc ggtgcgccgt   9060
15   cgattaggaa gccgcccaag ggcgacgagc aaccagattt tttcgttccg atgctctatg   9120
     acgtgggcac ccgcgatagt cgcagcatca tggacgtggc cgttttccgt ctgtcgaagc   9180
     gtgaccgacg agctggcgag gtgatccgct acgagcttcc agacgggcac gtagaggttt   9240
     ccgcagggcc ggccggcatg gccagtgtgt gggattacga cctggtactg atggcggttt   9300
     cccatctaac cgaatccatg aaccgatacc gggaagggaa gggagacaag cccggccgcg   9360
20   tgttccgtcc acacgttgcg gacgtactca agttctgccg gcgagccgat ggcggaaagc   9420
     agaaagacga cctggtagaa acctgcattc ggttaaacac cacgcacgtt gccatgcagc   9480
     gtacgaagaa ggccaagaac ggccgcctgg tgacggtatc cgagggtgaa gccttgatta   9540
     gccgctacaa gatcgtaaag agcgaaaccg ggcggccgga gtacatcgag atcgagctag   9600
     ctgattggat gtaccgcgag atcacagaag gcaagaaccc ggacgtgctg acggttcacc   9660
25   ccgattactt tttgatcgat cccggcatcg gccgttttct ctaccgcctg gcacgccgcg   9720
     ccgcaggcaa ggcagaagcc agatggttgt tcaagacgat ctacgaacgc agtggcagcg   9780
     ccggagagtt caagaagttc tgtttcaccg tgcgcaagct gatcgggtca aatgacctgc   9840
     cggagtacga tttgaaggag gaggcggggc aggctggccc gatcctagtc atgcgctacc   9900
     gcaacctgat cgagggcgaa gcatccgccg gttcctaatg tacggagcag atgctagggc   9960
30   aaattgccct agcaggggaa aaaggtcgaa aaggtctctt tcctgtggat agcacgtaca  10020
     ttgggaaccc aaagccgtac attgggaacc ggaacccgta cattgggaac ccaaagccgt  10080
     acattgggaa ccggtcacac atgtaagtga ctgatataaa agagaaaaaa ggcgattttt  10140
     ccgcctaaaa ctctttaaaa cttattaaaa ctcttaaaac ccgcctggcc tgtgcataac  10200
     tgtctggcca gcgcacagcc gaagagctgc aaaaagcgcc taccttcgg tcgctgcgct  10260
35   ccctacgccc cgccgcttcg cgtcggccta tcgcggccta tgcggtgtga ataccgcac  10320
     agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttcctcgct cactgactcg  10380
     ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg  10440
     ttatccacag aatcagggga taacgcagga agaacatgt gagcaaaagg ccagcaaaag  10500
     gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccctgac  10560
40   gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga  10620
     taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt  10680
     accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc  10740
     tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc  10800
     cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta  10860
45   agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat  10920
     gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca  10980
     gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct  11040
```

```
tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt    11100
acgcgcagaa aaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct     11160
cagtggaacg aaaactcacg ttaagggatt ttggtcatgc atgatatatc tcccaatttg    11220
tgtagggctt attatgcacg cttaaaaata ataaaagcag acttgacctg atagtttggc    11280
tgtgagcaat tatgtgctta gtgcatctaa cgcttgagtt aagccgcgcc gcgaagcggc    11340
gtcggcttga acgaatttct agctagacat tatttgccga ctaccttggt gatctcgcct    11400
ttcacgtagt ggacaaattc ttccaactga tctgcgcgcg aggccaagcg atcttcttct    11460
tgtccaagat aagcctgtct agcttcaagt atgacgggct gatactgggc cggcaggcgc    11520
tccattgccc agtcggcagc gacatccttc ggcgcgattt tgccggttac tgcgctgtac    11580
caaatgcggg acaacgtaag cactacattt cgctcatcgc cagcccagtc gggcggcgag    11640
ttccatagcg ttaaggtttc atttagcgcc tcaaatagat cctgttcagg aaccggatca    11700
aagagttcct ccgccgctgg acctaccaag caacgctat gttctcttgc ttttgtcagc     11760
aagatagcca gatcaatgtc gatcgtggct ggctcgaaga tacctgcaag aatgtcattg    11820
cgctgccatt ctccaaattg cagttcgcgc ttagctggat aacgccacgg aatgatgtcg    11880
tcgtgcacaa caatggtgac ttctacagcg cggagaatct cgctctctcc aggggaagcc    11940
gaagtttcca aaaggtcgtt gatcaaagct cgccgcgttg tttcatcaag ccttacggtc    12000
accgtaacca gcaaatcaat atcactgtgt ggcttcaggc cgccatccac tgcggagccg    12060
tacaaatgta cggccagcaa cgtcggttcg agatggcgct cgatgacgcc aactacctct    12120
gatagttgag tcgatacttc ggcgatcacc gcttccccca tgatgtttaa ctttgtttta    12180
gggcgactgc cctgctgcgt aacatcgttg ctgctccata acatcaaaca tcgacccacg    12240
gcgtaacgcg cttgctgctt ggatgcccga ggcatagact gtaccccaaa aaaacagtca    12300
taacaagcca tgaaaccgc cactgcgttc catg                                  12334
```

<210> 16
<211> 11385
<212> DNA
<213> STPT GUS ptx-P daao

<220>
<221> misc_feature
<222> (38)..(183)
<223> Agrobacterium Right Border

<220>
<221> terminator
<222> (384)..(639)
<223> complementary: nos terminator

<220>
<221> misc_feature
<222> (716)..(1822)
<223> complementary: coding for Rhodotorula gracilis DAAO

<220>
<221> promoter

```
     <222>   (1866)..(2728)
     <223>   complementary: ptxA promoter


     <220>
 5   <221>   terminator
     <222>   (2818)..(3022)
     <223>   complementary: 35S terminator


     <220>
10   <221>   misc_feature
     <222>   (3097)..(5094)
     <223>   complementary: coding for beta-glucuronidase (comprising intron)


     <220>
15   <221>   promoter
     <222>   (5148)..(6465)
     <223>   complementary: sTPT promoter


     <220>
20   <221>   misc_feature
     <222>   (6537)..(6753)
     <223>   Agrobacterium Left Border


     <400>   16
25   aatattcaaa caaacacata cagcgcgact tatcatggac atacaaatgg acgaacggat     60
     aaacctttc  acgccctttt  aaatatccga  ttattctaat aaacgctctt  ttctcttagg   120
     tttacccgcc aatatatcct gtcaaacact gatagtttaa actgaaggcg ggaaacgaca    180
     atcagatctg gtaccttgcg ccgggtaccc caaactgtct cacgacgttt tgaacccaga    240
     ttaccctgtt atccctagtc gagcggccgc cagtgtgatg gatatctgca gaattcgccc    300
30   ttttagatca gcacactggc ggccgttact agtggatcaa ttcactggcc gtcgttttac    360
     aacgactcag agcttgacag gaggcccgat ctagtaacat agatgacacc gcgcgcgata    420
     atttatccta gtttgcgcgc tatattttgt tttctatcgc gtattaaatg tataattgcg    480
     ggactctaat cataaaaacc catctcataa ataacgtcat gcattacatg ttaattatta    540
     catgcttaac gtaattcaac agaaattata tgataatcat cgcaagaccg gcaacaggat    600
35   tcaatcttaa gaaactttat tgccaaatgt ttgaacgatc ggggatcatc cgggtctgtg    660
     gcgggaactc cacgaaaata tccgaacgca gcaagatcta gagcttgggt cccgcctaca    720
     acttcgactc ccgcgccgcg ccgtggtacc gctggaacgc ctcgtcgacg agctgcgcga    780
     catcctccgc cgcgccccaa ctctgctggt atccgcact  cgagaagcca tacgcatgca    840
     caagcgtgac ctccttctcc ttcgccgctc gtgcgctgcc cctgccgagc gagaggggcg    900
40   actttgtccg gtcgagaggc aggacgatcc gttctgcctc aacgcggggt ccgcctcgtc    960
     gtgcaggtcg caagccgacg ttgtggcgga ggacctcgat gccttcgatc gttccgtcgc   1020
     tcgagatggt cgggtcgagg cgcaagcagt gcttgaggat ccgctggacc gtctctgggt   1080
     tgacagacaa gtcccagtct cccacgccgt acgtcccgcc gcagatgact tcgccacctg   1140
     gtcggggaat gatgtaggcg ggagaagcgg ggtcggacga gtccatcgtg catcgcttgc   1200
45   atggggactt gacgaggacg gtttgcccgc ggattggctc ggcggcttgg tcgtcgatgc   1260
     ccgcaatcga cttggcgcca agtcccgtag cgttgaccac caaatccgca ccgtcgaacg   1320
     cctgctcaag cgacgtaacg gtccgtctct caaacgtcgc gccgagcttc tgcagctctc   1380
```

```
ttgcaaggta ctggcagtac tttggtgcgt ggacggagag ggtgtcgtag gttacgccga   1440
tagcgccagg tggacattcg gaagatggga gggggcggta atttggcgtg atgtccttgt   1500
accagtgccc gagcaagccg tcttcgttct gcgcgaaccg cctcgtcccc ttgagccaca   1560
tggcatggcc cgtcgggacc aactcgaccc acttcttgaa agtcgattct cccatttttg   1620
cttgtcgagg accgtctgta agcgtcatga aaggcgtcca attcgcgcca gcccatggtg   1680
aagcgaaagt ctggctcgag acgtcctccg gcaagtcgcg cgcgagaata tgcacgctgt   1740
agcccttccg agcgaggatg agggcgctgc tcagaccgat aacgcctgat ccgaggacaa   1800
cgacgcgctt ctgcgagtgc atgggccctc gactagagtc gagatccgat atcgcccggg   1860
ctcgagtttc gaagatttta gtgtaatgtg tgtgctcact actatgaagc tttgcactta   1920
aaaaaataga atgagtgatg aggtttatat ggtgaaaaaa actatgaaat tttgatattt   1980
tgatatatct ttctcgtgag tcatattcac ggaccatgtt gcagcaaatt ggaattaaac   2040
tattcatttt ttatgttaaa tcattgattg attttttagtg ggcctcgtta catattcaag   2100
agttagaatg aattcaaaca aactaggcca gaaaaaagga tgtggggcca tttttttgtg   2160
tcttaagaat ttgtttattt ttttcatgga taaggggaat caatggaaaa agtttgatgt   2220
actagaggac atttttttaa catgtagtga caagtagtgc tattattcga cccgtgatga   2280
aaggggcaat cttaatcttt ttttcataaa tctgcacatg tgatgcttta attatgcttt   2340
agactttgtg ctaaactatt ggtaatttct ttttgtaatc gaatcaagta tcttttaaac   2400
tatgtatgaa atgtgtcatc ctaaaaacaa cattttgcta gttttagact ttgatgttta   2460
tatgcttaat ggaagaagca atatgttgat gtttattggg taaaagaaag ggacttgatt   2520
gagtatgtaa ttgacaacta tgattttata ttggatttga tattcctaac attaatttaa   2580
gtgtgtgggt ttcaaagcat gttatgctag tgattcttgt gtttgatgct tgaaaaatct   2640
acattcatcc ttgaatggag ggacaaactt tgaatgactt ttgaataggt gtaaaatcca   2700
atcctccctc agcttcacaa aaaattgcct cgaggtacct ggtagggata acagggtaat   2760
ctgggttcaa aacgtcgtga gacagtttgg tgcaggtcga aattcgagct cggtacccgg   2820
tcactggatt ttggttttag gaattagaaa ttttattgat agaagtattt tacaaataca   2880
aatacatact aagggtttct tatatgctca acacatgagc gaaaccctat aagaaccctа   2940
attcccttat ctgggaacta ctcacacatt attctggaga aaaatagaga gagatagatt   3000
tgtagagaga gactggtgat ttttgcgccg ggtaccgagc tcggtagcaa ttcccgaggc   3060
tgtagccgac gatggtgcgc caggagagtt gttgattcat tgtttgcctc cctgctgcgg   3120
ttttttcaccg aagttcatgc cagtccagcg ttttttgcagc agaaaagccg ccgacttcgg   3180
tttgcggtcg cgagtgaaga tcccttttctt gttaccgcca acgcgcaata tgccttgcga   3240
ggtcgcaaaa tcggcgaaat tccatacctg ttcaccgacg acggcgctga cgcgatcaaa   3300
gacgcggtga tacatatcca gccatgcaca ctgatactct tcactccaca tgtcggtgta   3360
cattgagtgc agcccggcta acgtatccac gccgtattcg gtgatgataa tcggctgatg   3420
cagtttctcc tgccaggcca gaagttcttt ttccagtacc ttctctgccg tttccaaatc   3480
gccgctttgg acataccatc cgtaataacg gttcaggcac agcacatcaa agagatcgct   3540
gatggtatcg gtgtgagcgt cgcagaacat tacattgacg caggtgatcg gacgcgtcgg   3600
gtcgagttta cgcgttgctt ccgccagtgg cgaaatattc ccgtgcactt gcggacgggt   3660
atccggttcg ttggcaatac tccacatcac cacgcttggg tggtttttgt cacgcgctat   3720
cagctctttta atcgcctgta agtgcgcttg ctgagtttcc ccgttgactg cctcttcgct   3780
gtacagttct ttcggcttgt tgcccgcttc gaaaccaatg cctaaagaga ggttaaagcc   3840
gacagcagca gtttcatcaa tcaccacgat gccatgttca tctgcccagt cgagcatctc   3900
ttcagcgtaa gggtaatgcg aggtacggta ggagttggcc ccaatccagt ccattaatgc   3960
gtggtcgtgc accatcagca cgttatcgaa tcctttgcca cgtaagtccg catcttcatg   4020
acgaccaaag ccagtaaagt agaacggttt gtggttaatc aggaactgtt cgcccttcac   4080
tgccactgac cggatgccga cgcgaagcgg gtagatatca cactctgtct ggcttttggc   4140
tgtgacgcac agttcataga gataaccttc acccggttgc cagaggtgcg gattcaccac   4200
```

```
       ttgcaaagtc ccgctagtgc cttgtccagt tgcaaccacc tgttgatccg catcacgcag    4260
       ttcaacgctg acatcaccat tggccaccac ctgccagtca acagacgcgt ggttacagtc    4320
       ttgcgcgaca tgcgtcacca cggtgatatc gtccacccag gtgttcggcg tggtgtagag    4380
       cattacgctg cgatggattc cggcatagtt aaagaaatca tggaagtaag actgcttttt    4440
 5     cttgccgttt tcgtcggtaa tcaccattcc cggcgggata gtctgccagt tcagttcgtt    4500
       gttcacacaa acggtgatac ctgcacatca acaaattttg gtcatatatt agaaaagtta    4560
       taaattaaaa tatacacact tataaactac agaaaagcaa ttgctatata ctacattctt    4620
       ttattttgaa aaaatatttt gaaatattat attactacta attaatgata attattatat    4680
       atatatcaaa ggtagaagca gaaacttacg tacacttttc ccggcaataa catacggcgt    4740
10     gacatcggct tcaaatggcg tatagccgcc ctgatgctcc atcacttcct gattattgac    4800
       ccacactttg ccgtaatgag tgaccgcatc gaaacgcagc acgatacgct ggcctgccca    4860
       acctttcggt ataaagactt cgcgctgata ccagacgttg cccgcataat tacgaatatc    4920
       tgcatcggcg aactgatcgt taaaactgcc tggcacagca attgcccggc tttcttgtaa    4980
       cgcgctttcc caccaacgct gaccaattcc acagttttcg cgatccagac tgaatgccca    5040
15     caggccgtcg agttttttga tttcacgggt tggggtttct acaggacgta ccatggtcga    5100
       tcgactctag actagtggat ccgatatcgc ccgggctcga ctctagatga aatcgaaatt    5160
       cagagttttg atagtgagag caaagaggga cggacttatg aggatttcga gtatttcaag    5220
       agatggtact tgttgatcgg acggctacga tgatctcgat ttggttaatc cagtatctcg    5280
       cggtgtatgg agttatggta gggttaatgg tcaatttcat ctaacggtag agaatgatgt    5340
20     aattagataa gaatcttgag atactggttt agattggatg agtgtagggt ccatcttatc    5400
       ttgataagtg gatggttttt agagacacag tgaatattag ccaatcgaag ttccatatca    5460
       ccatcatcat ctgtataatt ttgtttttttt ggaagataat aatgattgaa attttggtag    5520
       atttttatttt tcattattta ccttgtatgt tgagtggtct tcaaattatt gaacgtgaca    5580
       gattcacaag aaagtagatt ttttataaat gaaattttac ttattttaaa ggtatctcta    5640
.25    tttaatttct tttgtttatg gttgtctgtc agcatttgac ttgcagtttc atgctcatag    5700
       tcatatacgt tattctaggc ttttttgaat atcttattac ttttttcgta atacaatttt    5760
       ataattttat caaagttata caactataac taaaattagg gtttttctaca aaacaaaaaa    5820
       atcttctaat ttttttttgtt gtagccagtt tactcgtaag ttacaaaaaa atacaaatga    5880
       acccacatgt attatgcgtt taactaggat taccatgtac tttcatgtac tcaattcacc    5940
30     ctatactctt tttttttttt tttctagttc cacccaatct ataaaattct gtccatttga    6000
       ccaaattcaa ttaatttctg taattgcgat ttaaaattaa tattacatgt tcactatttc    6060
       tcgatttgag ggaacccgag tttaaatatg ataaaaatgt tgacccatca ctacaaatat    6120
       gttatagttt atacttaata gtggtgtttt tggggataat tgatgaatta agtaaacatg    6180
       attcttctta tgaagttgat tgagtgatta ttgtatgtaa acctatgtga ttgatgttat    6240
35     tggttgattg agtgattatt gtattagtat gtaagcaaag atgattgttc ttatgaggta    6300
       atttgttact cattcatcct tttgcatatg agaaattgtg ttagcgtacg caaaacaata    6360
       gagaacataa aagatatgtg tatttatttta aggtgacttt tgttaatgat attgtagtat    6420
       ctatacattt atatataact tgttgaattt gagtataagc tatcaggatc cggggggatcc    6480
       tctagagtcg aggtaccgag ctcgaattca ctggccgtcg ttttacaacg actcagtact    6540
40     gcttggtaat aattgtcatt agattgtttt tatgcataga tgcactcgaa atcagccaat    6600
       tttagacaag tatcaaacgg atgttaattc agtacattaa agacgtccgc aatgtgttat    6660
       taagttgtct aagcgtcaat ttgtttacac cacaatatat cctgccacca gccagccaac    6720
       agctccccga ccggcagctc ggcacaaaat caccacgcgt ctaaaaggt gatgtgtatt    6780
       tgagtaaaac agcttgcgtc atgcggtcgc tgcgtatatg atgcgatgag taaataaaca    6840
45     aatacgcaag gggaacgcat gaaggttatc gctgtactta accagaaagg cgggtcaggc    6900
       aagacgacca tcgcaaccca tctagcccgc gccctgcaac tcgccggggc cgatgttctg    6960
       ttagtcgatt ccgatcccca gggcagtgcc cgcgattggg cggccgtgcg gaagatcaa    7020
```

```
     ccgctaaccg ttgtcggcat cgaccgcccg acgattgacc gcgacgtgaa ggccatcggc   7080
     cggcgcgact tcgtagtgat cgacggagcg ccccaggcgg cggacttggc tgtgtccgcg   7140
     atcaaggcag ccgacttcgt gctgattccg gtgcagccaa gcccttacga catatgggcc   7200
     accgccgacc tggtggagct ggttaagcag cgcattgagg tcacggatgg aaggctacaa   7260
 5   gcggcctttg tcgtgtcgcg ggcgatcaaa ggcacgcgca tcggcggtga ggttgccgag   7320
     gcgctggccg ggtacgagct gcccattctt gagtcccgta tcacgcagcg cgtgagctac   7380
     ccaggcactg ccgccgccgg cacaaccgtt cttgaatcag aacccgaggg cgacgctgcc   7440
     cgcgaggtcc aggcgctggc cgctgaaatt aaatcaaaac tcatttgagt taatgaggta   7500
     aagagaaaat gagcaaaagc acaaacacgc taagtgccgg ccgtccgagc gcacgcagca   7560
10   gcaaggctgc aacgttggcc agcctggcag acacgccagc catgaagcgg gtcaactttc   7620
     agttgccggc ggaggatcac accaagctga agatgtacgc ggtacgccaa ggcaagacca   7680
     ttaccgagct gctatctgaa tacatcgcgc agctaccaga gtaaatgagc aaatgaataa   7740
     atgagtagat gaattttagc ggctaaagga ggcggcatgg aaaatcaaga caaccaggc   7800
     accgacgccg tggaatgccc catgtgtgga ggaacgggcg gttggccagg cgtaagcggc   7860
15   tgggttgtct gccggccctg caatggcact ggaacccca agcccgagga atcggcgtga   7920
     gcggtcgcaa accatccggc ccggtacaaa tcggcgcggc gctgggtgat gacctggtgg   7980
     agaagttgaa ggccgcgcag ccgcccagc ggcaacgcat cgaggcagaa gacgccccgg   8040
     tgaatcgtgg caaggggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc   8100
     cggtgcgccg tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc   8160
20   gatgctctat gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg   8220
     tctgtcgaag cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca   8280
     cgtagaggtt tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact   8340
     gatggcggtt tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa   8400
     gcccggccgc gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga   8460
25   tggcggaaag cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt   8520
     tgccatgcag cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga   8580
     agccttgatt agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga   8640
     gatcgagcta gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct   8700
     gacggttcac cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct   8760
30   ggcacgccgc gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg   8820
     cagtggcagc gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc   8880
     aaatgacctg ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt   8940
     catgcgctac cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca   9000
     gatgctaggg caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga   9060
35   tagcacgtac attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa   9120
     cccaaagccg tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa   9180
     aggcgatttt tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc   9240
     ctgtgcataa ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg   9300
     gtcgctgcgc tccctacgcc ccgccgcttc gcgtcggcct atcgcggcct atgcggtgtg   9360
40   aaataccgca cagatgcgta aggagaaaat accgcatcag cgctcttcc gcttcctcgc   9420
     tcactgactc gctgcgctcg tcgttcggc tgcggcgagc ggtatcagct cactcaaagg   9480
     cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag   9540
     gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc   9600
     gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga acccgacag   9660
45   gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga   9720
     ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg cgctttctc   9780
     atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg   9840
```

```
tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt    9900
ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca    9960
gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca   10020
ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag   10080
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca   10140
agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg   10200
ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg catgatatat   10260
ctcccaattt gtgtagggct tattatgcac gcttaaaaat aataaaagca gacttgacct   10320
gatagtttgg ctgtgagcaa ttatgtgctt agtgcatcta acgcttgagt taagccgcgc   10380
cgcgaagcgg cgtcggcttg aacgaatttc tagctagaca ttatttgccg actaccttgg   10440
tgatctcgcc tttcacgtag tggacaaatt cttccaactg atctgcgcgc gaggccaagc   10500
gatcttcttc ttgtccaaga taagcctgtc tagcttcaag tatgacgggc tgatactggg   10560
ccggcaggcg ctccattgcc cagtcggcag cgacatcctt cggcgcgatt ttgccggtta   10620
ctgcgctgta ccaaatgcgg gacaacgtaa gcactacatt tcgctcatcg ccagcccagt   10680
cgggcggcga gttccatagc gttaaggttt catttagcgc ctcaaataga tcctgttcag   10740
gaaccggatc aaagagttcc tccgccgctg gacctaccaa ggcaacgcta tgttctcttg   10800
cttttgtcag caagatagcc agatcaatgt cgatcgtggc tggctcgaag atacctgcaa   10860
gaatgtcatt gcgctgccat tctccaaatt gcagttcgcg cttagctgga taacgccacg   10920
gaatgatgtc gtcgtgcaca acaatggtga cttctacagc gcggagaatc tcgctctctc   10980
cagggaagc cgaagtttcc aaaaggtcgt tgatcaaagc tcgccgcgtt gtttcatcaa   11040
gccttacggt caccgtaacc agcaaatcaa tatcactgtg tggcttcagg ccgccatcca   11100
ctgcggagcc gtacaaatgt acggccagca acgtcggttc gagatggcgc tcgatgacgc   11160
caactacctc tgatagttga gtcgatactt cggcgatcac cgcttccccc atgatgttta   11220
actttgtttt agggcgactg ccctgctgcg taacatcgtt gctgctccat aacatcaaac   11280
atcgacccac ggcgtaacgc gcttgctgct tggatgcccg aggcatagac tgtaccccaa   11340
aaaaacagtc ataacaagcc atgaaaaccg ccactgcgtt ccatg                   11385
```